Annals of Mathematics
Vol. 38, No. 4, October, 1937

THE REPRESENTATION OF A DEFINITE QUADRATIC FORM AS A SUM OF TWO OTHERS

By L. J. Mordell

(Received December 28, 1936)

The quadratic forms dealt with in this paper are all of the classic type

$$f(x) = \sum_{r,s=1}^{n} a_{rs}x_r x_s \qquad (a_{rs} = a_{sr}) \tag{1}$$

with integer coefficients a_{rs} and determinant

$$A = \| a_{rs} \| \qquad (r, s = 1, 2, \cdots, n). \tag{2}$$

The form is *positive definite* if for all real x except simultaneous zeros, $f(x) > 0$. This occurs if and only if A and the successive determinants A_{11}, $A_{11,22}$, $\cdots$ obtained from A by omitting the first row and column, the first two rows and columns, etc., are all positive.

The form is called *semi-definite* if $f(x) \geqq 0$ for all real x and if also $f(x) = 0$ for a set of values of the x's not all zero, but $f(x) \neq 0$ identically. Now $A = 0$ and it is known that every semi-definite form can be derived from a positive definite form of fewer variables on replacing them by linear functions of x_1, $x_2, \cdots x_n$ with integer coefficients. For since $f(x)$ represents zero, it is obviously transformable by a linear substitution with integer coefficients and determinant unity into

$$2y_1(b_{12}y_2 + b_{13}y_3 + \cdots) + g(y_2, y_3, \cdots y_n),$$

say. Since $f(x)$ is semi-definite, $b_{12} = b_{13} = \cdots = 0$, as otherwise $f(x)$ might assume negative values. A similar argument now applies to $g(y_2, y_3, \cdots y_n)$ etc.

It is clear that $f(x)$ is certainly semi-definite if $A = 0$, $A_{11} > 0$, $A_{11,22} > 0$, etc.

Some years ago I proposed the following problem—Is a decomposition

$$f(x) = \sum_{r=1}^{N} (L_r(x))^2 \tag{3}$$

possible, where the L's are linear forms with integer coefficients and N is independent of $f(x)$?—and suggested[1] that such a decomposition exists with $N = n + 3$. This would be the best possible value of N, for it is easy to see that the form

[1] Mordell, Mathematische Zeitschrift, 35 (1932) 1–15. The date of reception should be May 31, 1930 not 1931.

$$f(x) = \sum_{r=1}^{n-1} x_r^2 + 7x_n^2$$

cannot be decomposed into (3) with a smaller value of N than $n + 3$.

I proved this[2] for $n = 2$ and Mr. Chao Ko[3] has proved it for $n = 3, 4$. His method of proof differs from mine. It depends upon expressing $f(x)$ as part of a definite form in $n + 3$ variables of determinant unity and using the theorem that there is only one class of positive forms for $A = 1$ and $n \leqq 7$. It was not known whether (3) holds for $N = n + 3$ and $n > 4$. He has found since at the same time as myself that there is also only one class of positive properly primitive forms when $n = 8$ and $A = 1$. Hence the hypothesis holds for $n = 5$. I may add that the results of this paper prove that the hypothesis is false for $n = 6$.

The earlier work on this subject has suggested to Dr. Erdös the following problem which he and Mr. Ko have proposed to me. Can $f(x)$ be expressed as a sum of two definite quadratic forms with integer coefficients? The case $f(x) = x_1^2 + (x_1 + x_2)^2$ suggests that it is desirable to modify the question and to propose the

Problem. If $f(x) = \sum_{r,s=1}^{n} a_{rs}x_r x_s\ (a_{rs} = a_{sr})$ is a definite quadratic form with integer coefficients $a_{r,s}$ is a decomposition

$$f(x) = f_1(x) + f_2(x) \tag{4}$$

always possible where $f_1(x)$, $f_2(x)$ are positive definite or semi-definite forms with integer coefficients $a_{r,s}^{(1)}$, $a_{r,s}^{(2)}$?

The problem seems simple enough and one might reasonably expect that the solution would not involve any advanced theory. It turns out however, to be a very deep one in the arithmetical theory of quadratic forms in several variables. By such means, I prove that (4) is always possible when $n \leqq 5$, and that when $n > 5$ is given, (4) is possible when $A > \lambda_n$ where λ_n depends only on n, e.g. $\lambda_n = (4/3)^{\frac{1}{2}n(n-1)}$ suffices. Further, I give reasons which suggest that the decomposition (4) may be impossible if $n > 5$ and A is too small (but not so small that the form represents 1), e.g. $A < \mu_n$ where μ_n is a function of n whose existence is easily demonstrated but of which the best possible value is not easily found. Following on this, Mr. Chao Ko has proved that (4) is impossible for at least one form of 6 or 7 variables or for a form in 8 variables with determinant unity. I show, however, that (4) is impossible for the form in 6 variables of determinant 3, namely,

$$g(x) = \sum_{r=1}^{6} x_r^2 + \left(\sum_{1}^{6} x_r\right)^2 - 2x_1x_2 - 2x_2x_6. \tag{5}$$

The result for $n \leqq 5$ is of course included in Mr. Chao Ko's result but the proof is simpler than his as less is proved. The example (5) shows that no result of

---

[2] Mordell, Quarterly Journal of Mathematics, 1 (1930), 276–288.

[3] Chao Ko, Quarterly Journal of Mathematics, 8 (1937), 81–98.

the form (3) holds, i.e. if L is a linear form with integer coefficients, $g(x) - L^2$ is an *indefinite form.*

Take first the case $n = 2$. It suffices to replace $f(x)$ by an equivalent reduced form for which $a_{12} \geqq 0$, $a_{22} \geqq 2a_{12}$, $a_{11} \geqq 2a_{12}$. Then

$$f(x) = (a_{11} - a_{12})x_1^2 + a_{12}(x_1 + x_2)^2 + (a_{22} - a_{12})x_2^2 .$$

Since $a_{11} \geqq a_{12}$, $a_{22} \geqq a_{12}$, a decomposition (4) is always possible except when $a_{11} = a_{12} = a_{22} = 1$, really corresponding to the excluded semi-definite form $(x_1 + x_2)^2$ of determinant zero; or again when $a_{11} = a_{12} = 0$, $a_{22} = 1$ etc.

Take next $n = 3$. Then by Selling's[4] method of reduction, there exists for $f(x)$ a reduced form given by

$$f(x) = \rho_1 x_1^2 + \rho_2 x_2^2 + \rho_3 x_3^2 + \rho_4(x_2 - x_3)^2 + \rho_5(x_3 - x_1)^2 + \rho_6(x_1 - x_2)^2,$$

where no ρ is negative. On comparing coefficients, clearly all the ρ's are integers. Here again a decomposition (4) is always possible. The case when all the ρ's are zero except one which is unity gives $A = 0$.

For $n > 3$, it is also true that any definite quadratic form is equivalent to an $f(x)$ given by

$$f(x) = \sum_{t=1}^{m} \rho_t L_t^2 \tag{6}$$

where the L_t are linear forms whose coefficients are integers and depend only on n (as does also m) and not on $f(x)$, and the ρ's are non-negative rational numbers.

This result, however, is part of Voronoi's[5] theory of quadratic forms. It is not obvious or certain that the ρ's are integers when the coefficients of $f(x)$ are integers. A decomposition (4) is obviously possible if at least one of the ρ's is not less than unity, e.g. if $\rho_1 > 1$, on writing $\rho_1 = [\rho_1] + (\rho_1 - [\rho_1])$ where $[\rho_1]$ denotes the integer part of ρ_1. It may, however, be impossible if all the ρ's are less than unity, and this implies from (6) for the determinant A of $f(x)$ that $A \leqq \delta_n$, where the value of δ_n depends only on n but is not easily calculated. Hence it is suggested that (4) will be possible for $A > \mu_n$, where μ_n depends only on n.

An explicit expression for a λ_n such that (4) holds for $A > \lambda_n$ is given by the λ_n of the following argument. A well known result by Hermite[6] states that $f(x)$ is equivalent to a reduced form for which

$$a_{11} \leqq a_{22} \cdots \leqq a_{nn}, \qquad a_{11} \leqq (\lambda_n A)^{1/n}, \tag{7}$$

where λ_n is a number, e.g. $(4/3)^{\frac{1}{2}n(n-1)}$ depending only on n. The determination

[4] See Bachmann, *Die Arithmetik der Quadratischer Formen*, 2 (1923), 414–418.

[5] Bachmann, 292–318.

[6] Bachmann, 250–255.

of the best possible value of λ_n is a question of great difficulty. It has been known for a long time[7] that

$$\lambda_2 = 4/3, \qquad \lambda_3 = 2, \qquad \lambda_4 = 2, \qquad \lambda_5 = 8,$$

and fairly recently it has been shown that

$$\lambda_6 = 64/3, \qquad \lambda_7 = 64, \qquad \lambda_8 = 256.$$

The proof for λ_6 has been given both by Hofreiter[8] and Blichfeldt[9] and for λ_7, λ_8 by Blichfeldt.

Denote by A_{rs} the cofactors of the a_{rs} in A, e.g.

$$A_{11} = \| a_{rs} \| \qquad (r, s = 2, 3 \cdots n),$$

and by

$$F(x) = \sum_{r,s=1}^{n} A_{rs} x_r x_s,$$

the adjoint form of $f(x)$. Then the determinant of $F(x)$ is A^{n-1}.

Suppose now that $f(x)$ is replaced by an equivalent form for which $F(x)$ is reduced, so that corresponding to (7)

$$A_{11} \leqq (\lambda_n A^{n-1})^{1/n}. \tag{8}$$

I now prove that if $A \geqq \lambda_n$, a decomposition (4) with $f_2(x) = x_1^2$, i.e.

$$f(x) = x_1^2 + f_1(x) \tag{9}$$

is possible. For since $A_{11} > 0$, $A_{11,22} > 0$ etc., $f(x) - x_1^2$ is neither definite nor semi-definite if and only if its determinant

$$\left\| \begin{matrix} a_{11} - 1, & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \cdots & \cdots & \cdots \end{matrix} \right\| < 0, \text{ i.e.}$$

$$A < A_{11}. \tag{10}$$

Then of course, since $A_{11} \leqq A_{22} \leqq \cdots$, none of $f(x) - x_r^2$ is definite or semi-definite. From (10) and (8),

$$A < (\lambda_n A^{n-1})^{1/n}, \quad \text{i.e.} \quad A < \lambda_n.$$

Hence (9) is possible if $A \geqq \lambda_n$. Clearly $f_1(x)$ is semi-definite if $A = A_{11} = \lambda_n$.

Consider now the cases when $A < \lambda_n$. We suppose that $f(x)$ is reduced by (7), i.e. not $F(x)$ as before. Then

$$a_{11} < (\lambda_n^2)^{1/n}.$$

[7] Bachmann, 327–328.

[8] Hofreiter, Monatshefte für Mathematik und Physik, 40 (1933), 129–152.

[9] Blichfeldt, Mathematische Zeitschrift, 39 (1934) 1–15.

Take first $n = 2, 3, 4$. Then $a_{11} < 2$ and so $a_{11} = 1$. Hence

$$(11)\qquad \begin{aligned} f(x) &= x_1^2 + 2a_{12}x_1x_2 + \cdots \\ &= (x_1 + a_{12}x_2 + \cdots)^2 + f_1(x_2, x_3, \cdots), \end{aligned}$$

where $f_1(x_2, x_3, \cdots)$ is not an indefinite form.

Hence (9) is possible.

Take next the case when $n = 5$, so that we must examine the forms $f(x)$ with determinant $A \leqq 7$. If $A < 3$, there is a reduced form equivalent to $f(x)$ with $a_{11} \leqq 24^{\frac{1}{5}}$, i.e. $a_{11} = 1$, and a glance at (11) shows that the decomposition is possible. Hence we must examine the cases $A = 4, 5, 6, 7$ and a reduced $F(x)$ wherein

$$A < A_{11} \leqq (8A^4)^{\frac{1}{5}}.$$

But for these A, this would give $A < (8A^4)^{\frac{1}{5}} < A + 1$, and so $A = A_{11}$. Hence there exists a decomposition (9) where $f_1(x)$ is semi-definite. This proves (4) for $n = 5$.

Hence for $n \leqq 5$, (4) also holds for semi-definite forms $f(x)$ except when $f(x)$ is equivalent to x_1^2.

We take now the case $n = 6$, so that we must discuss the A for which

$$A < A_{11} \leqq (64A^5/3)^{\frac{1}{6}}$$

is possible. This requires $A < 14$.

The cases $A = 1, 2$ lead to $a_{11} = 1$ and (4) is possible. Hence we must consider

$$3 \leqq A \leqq 14.$$

Now forms with $A = 3$ exist such that for integers x, $f(x) = (\lambda_n A)^{\frac{1}{n}}$, and are known in connection with the *extreme* forms of Korkine[10] and Zolotereff, Hofreiter[8] and Blichfeldt.[9] In particular (5) is one. A decomposition (4) is impossible for (5). For if it were possible, there would exist one with f_2 equal to L^2. Thus either f_1 or f_2 must have a coefficient of x_r^2 equal to unity and then a decomposition with $f_1 = L_1^2$ is proved; or f_1, say, is definite or semi-definite, the coefficient of an x_r^2 being zero. But the decomposition with L_1^2 has been proved true for forms of not more than 5 variables. A similar argument was used by Mr. Chao Ko for his *Gegenbeispiel*.

To prove the impossibility of (4) for $g(x)$ write it as

$$g(x) = 2(x_1 + \tfrac{1}{2}(x_3 + x_4 + x_5 + x_6))^2 + 2(x_2 + \tfrac{1}{2}(x_3 + x_4 + x_5))^2 + \sum_{r=3,4,5} (x_r + \tfrac{1}{2}x_6)^2 + \tfrac{3}{4}x_6^2.$$

[10] Korkine and Zolotareff, Mathematische Annalen, 6 (1873), 366–389. Also Bachmann, 318–328.

Now the adjoint form of $\sum_{r=1}^{6} c_r(b_{r1}x_1 + b_{r2}x_2 + \cdots)^2$ is given by $(\prod_{r=1}^{6} c_r)$ $\sum_{r=1}^{6} X_r^2/c_r$,

where $\quad x_1 = b_{11}X_1 + b_{21}X_2 + \cdots$ etc.

Here $\quad x_1 = X_1, \quad x_2 = X_2,$

$$x_r = \tfrac{1}{2}X_1 + \tfrac{1}{2}X_2 + X_r \quad (r = 3, 4, 5),$$

$$x_6 = \tfrac{1}{2}X_1 + \tfrac{1}{2}(X_3 + X_4 + X_5) + X_6 .$$

Hence $\quad X_1 = x_1, \quad X_2 = x_2, \quad X_r = x_r - \tfrac{1}{2}(x_1 + x_2), \quad (r = 3, 4, 5),$

$$X_6 = x_6 - \tfrac{1}{2}x_1 - \tfrac{1}{2}(x_3 + x_4 + x_5) + \tfrac{3}{4}(x_1 + x_2).$$

Hence the adjoint form $F(x)$ is $\frac{3}{2}\, G(x)$, where

$$G(x) = x_1^2 + x_2^2 + 2 \sum_{r=3,4,5} (x_r - \tfrac{1}{2}(x_1 + x_2))^2$$

$$+ \tfrac{8}{3}(x_6 - \tfrac{1}{2}x_1 - \tfrac{1}{2}(x_3 + x_4 + x_5) + \tfrac{3}{4}(x_1 + x_2))^2.$$

It is now shown that for all integers x except simultaneous zeros, $G(x) \geqq 8/3$, $F(x) \geqq 4$, i.e. $A - A_{11} < 0$ and so $f(x) - L^2$ is an indefinite form.

Consider then the integers x for which $G(x) < 8/3$. Clearly $x_1^2 \leqq 1$, $x_2^2 \leqq 1$. Since $G(x)$ is unaltered by changing the signs of all the x's, we need discuss only the following cases:

(I)	$x_1 = 1,$	$x_2 = 1,$
(II)	$x_1 = 1,$	$x_2 = -1,$
(III)	$x_1 = 0,$	$x_2 = 0,$
(IV)	$x_1 = 1,$	$x_2 = 0,$
(V)	$x_1 = 0,$	$x_2 = 1.$

For (I), $G(x) = 2 + 2(x_3 - 1)^2 + \cdots,$
and so $x_3 = x_4 = x_5 = 1$. Then

$$G(x) = 2 + \tfrac{8}{3}(x_6 - \tfrac{1}{2})^2 \geqq 2 + \tfrac{2}{3} = \tfrac{8}{3}$$

For (II), $G(x) = 2 + 2x_3^2 + \cdots,$
and so $x_3 = x_4 = x_5 = 0$. Then

$$G(x) = 2 + \tfrac{8}{3}(x_6 - \tfrac{1}{2})^2 \geqq \tfrac{8}{3}$$

For (III), $G(x) = 2x_3^2 + 2x_4^2 + 2x_5^2 + \frac{8}{3}(x_6 - \frac{1}{2}(x_3 + x_4 + x_5))^2$.
Clearly two of x_3, x_4, x_5, say x_4, x_5 must be zero. Then if $x_3 = \pm 1$, $G(x) \geqq 2 + \frac{8}{3}(\frac{1}{4}) = \frac{8}{3}$. If $x_3 = 0$, $G(x) = \frac{8}{3}x_6^2 \geqq \frac{8}{3}$ unless $x_6 = 0$.
For (IV), $G(x) = 1 + 2 \sum_{r=3;4,5} (x_r - \frac{1}{2})^2 + \frac{8}{3}(x_6 + \frac{1}{4} - \frac{1}{2}(x_3 + x_4 + x_5))^2$

$$\geqq 1 + \tfrac{3}{2} + \tfrac{8}{3}(\tfrac{1}{4})^2 = \tfrac{8}{3}.$$

For (V), $G(x) \geqq 1 + \frac{3}{2} + \frac{8}{3}(x_6 + \frac{3}{4} - \frac{1}{2}(x_3 + x_4 + x_5))^2$

$$\geqq 1 + \tfrac{3}{2} + \tfrac{8}{3}(\tfrac{1}{4})^2 = \tfrac{8}{3}.$$

Hence $G(x) < \frac{8}{3}$ only when all the x's are zero.

The discussion of the other A for which (4) may be impossible seems a very interesting problem. It should be noted that if $n \geqq 6$ and a decomposition (4) exists wherein $f_1(x)$ is a quadratic form in not more than five variables, our results show that a decomposition (4) is also possible with $f_1(x) = (L_1(x))^2$. Hence further results cannot be expected unless we take $f_1(x)$ to be a form in at least six variables e.g. (5), which does not itself admit of a decomposition such as (4). The conditions, however, that $f(x) - f_1(x)$ should be definite or semi-definite, are not sufficiently simple to suggest that it will be easy to find results of interest.

There are, however, the *extreme* forms of Korkine and Zolotareff in seven and eight variables, namely

$$h(x) = \sum_1^7 x_r^2 + \left(\sum_1^7 x_r\right)^2 - 2x_1x_2 - 2x_2x_7$$

of determinant 2, and

$$k(x) = \sum_1^8 x_r^2 + \left(\sum_1^8 x_r\right)^2 - 2x_1x_2 - 2x_2x_8$$

of determinant[11] 1, which are easily shown to be undecomposable by (4).

Thus since $g(x)$ results on taking $x_6 = 0$ in $h(x)$ and then replacing x_7 by x_6, it is clear that a decomposition (4) cannot exist for $h(x)$ with $f_1(x) = (L_1(x))^2$, for then there would exist a decomposition for $g(x)$. Hence also there cannot exist a decomposition of $h(x)$ with terms x_1^2, x_2^2, $\cdots$ etc. occuring in either $f_1(x)$ or $f_2(x)$. But if a term say $2x_3^2$ occurs in $f_1(x)$, then $f_1(x)$ must also contain all the terms of $h(x)$ involving x_3, e.g. $2x_1x_3$, etc. For if a term in x_1x_3 occurred in $f_2(x)$, $f_2(x)$ could assume negative values by choice of x_3. Hence by the same argument, $f_1(x)$ must include all terms involving x_1, x_2, x_4, x_5, x_6, x_7. This proves that $f_2(x) = 0$.

A similar argument applies to the form $k(x)$.

UNIVERSITY, MANCHESTER, ENGLAND.

[11] I have shown that $k(x)$ and $\sum_{r=1}^8 x_r^2$ are the only positive classes of forms of determinant 1 in eight variables. The proof will appear in the Journal de Mathématiques.

ANNALS OF MATHEMATICS
Vol. 38, No. 4, October, 1937

ON THE CONVERGENCE PROPERTIES OF LAGRANGE INTERPOLATION BASED ON THE ZEROS OF ORTHOGONAL TCHEBYCHEFF POLYNOMIALS[1]

BY J. SHOHAT

(Received January 29, 1937)

Introduction. Statement of results. Let $f(x)$ be finite at every point of a certain interval (a, b). Consider Lagrange interpolation formula (LIF) using for abscissas the zeros

$$(1)\qquad x_{1,n} < x_{2,n} < \cdots < x_{n,n} \qquad (a < x_{1,n};\ x_{n,n} < b)$$

of the orthonormal Tchebycheff polynomials[2] (OP)

$$(2)\qquad \varphi_n(x; d\psi) \equiv \varphi_n(x) \equiv \varphi_n(x; a, b; d\psi) = a_n[x^n - S_n x^{n-1} + \cdots] \qquad (n = 1, 2, \cdots);$$
$$a_n \equiv a_n(d\psi) > 0, \qquad S_n \equiv S_n(d\psi).$$

$$(3)\qquad f(x) \approx \sum_{i=1}^{n} \frac{\varphi_n(x)}{(x - x_i)\,\varphi_n'(x_i)} f(x_i) \qquad (x_i \equiv x_{i,n}).$$

We rewrite (3) as

$$(4)\qquad f(x) = \sum_{i=1}^{n} l_i(x) f(x_i) + \rho_n(f) \equiv L_n(f) + \rho_n,$$

where

$$(5)\qquad l_i(x) = \frac{\varphi_n(x)}{(x - x_i)\,\varphi_n'(x_i)}, \qquad L_n(f) = \sum_{i=1}^{n} l_i(x) f(x_i).$$

$l_i(x)$ are called the "basic polynomials" of LIF under consideration. Hereafter, (a, b) is supposed to be finite, reduced, if we wish, to $(-1, 1)$; $f(x)$ is assumed to be bounded on (a, b).

It is well known (Faber, S. Bernstein) that whatever may be the choice of abscissas in LIF, we cannot get convergence—$\lim_{n\to\infty} \rho_n(f) = 0$—for *all* continuous $f(x)$. On the other hand, if the abscissas are the zeros $x_{i,n}$ of OP, as in (3), and if we assume (here and hereafter) the existence of $\int_a^b f(x)\, d\psi$, then

$$(6)\qquad \lim_{n\to\infty} \int_a^b \rho_n d\psi \equiv \lim_{n\to\infty} \int_a^b [f(x) - L_n(f)]\, d\psi = 0,$$

[1] Presented to the American Mathematical Society, March 26, 1937.

[2] The notations herein employed are the same as in my Monograph: *Théorie générale des polynomes orthogonaux de Tchebicheff*, Mémorial des sciences mathématiques, Fasc. 66, 1934.

and this is one of the classical results in the theory of OP. The question then arises as to the existence of some other convergence properties of the interpolation polynomial $L_n(f)$ in (4). The special case of $d\psi(x) = p(x)\,dx$, with $p(x) \geqq M > 0$ in $(-1, 1)$, and more particularly, $d\psi(x) = dx/(1 - x^2)^{\frac{1}{2}}$, has been studied recently by Erdös, Turán and Feldheim.[3] The main results here are the following:

$$\lim_{n\to\infty} \int_{-1}^{1} [f(x) - L_n(f)]^2 p(x)\,dx = 0 \tag{7.1}$$

$$\lim_{n\to\infty} \int_{-1}^{1} |f(x) - L_n(f)|^r \frac{dx}{\sqrt{(1 - x^2)}} = 0,^4 \qquad r > 0, \text{ arbitrary,} \tag{7.2}$$

where $f(x)$ is assumed bounded and R-integrable in $(-1, 1)$.

The main object of the present paper is to prove the following general

THEOREM A. *Let $f(x)$ be finite at every point of the finite interval (a, b) and such that $\int_a^b f^2(x)\,dx$ exists. Then*

$$\lim_{n\to\infty} \int_a^b |f(x) - L_n(f)|^r d\psi = 0, \qquad 0 < r \leqq 2, \tag{7}$$

if $\{\varphi_n(x)\}$ in (3) represents any system of OP *corresponding to (a, b).*

Our paper has several points in common with the third paper cited above. But the method of attack is different. It may be characterized summarily as "from mechanical quadratures to convergence," that is to say, the convergence properties of LIF (3) are derived from the classical properties of the mechanical quadratures formula (MQF) generated by (4):

$$\int_a^b f(x)dx = \sum_{i=1}^{n} H_i f(x_i) + R_n(f), \tag{8}$$

---

[3] 1) Ervin Feldheim, *Sur l'orthogonalité des fonctions fondamentales de l'interpolation de Lagrange*, Comptes Rendus, v. 203, 1936, 650–652;

2) P. Erdös et Ervin Feldheim, *Sur le mode de convergence pourl' interpolation de Lagrange*, ibid., 913–915;

3) P. Erdös and P. Turán, *On Interpolation I. Quadrature—and mean—convergence for Lagrange Interpolation*, these Annals, v. 38, 1937, 142–155.

[4] Let formula (7.2) be established for r = any positive even integer $2k$. The extension to any positive r is attained by letting in the Minkowsky-Hölder inequality

$$\left| \int_a^b f_1(x) f_2(x)\,d\psi(x) \right| \leqq \left[\int_a^b |f_1(x)|^s\,d\psi \right]^{1/s} \cdot \left[\int_a^b |f_2(x)|^{s/(s-1)}\,d\psi \right]^{(s-1)/s} \quad (s > 1):$$

$$f_1(x) = 1, \qquad f_2(x) = |\rho_n|^r, \qquad s = 2k/(2k - r) \qquad (2k > r).$$

We thus get:

$$\int_a^b |\rho_n|^r\,d\psi \leqq \left[\int_a^b d\psi \right]^{1/s} \cdot \left[\int_a^b |\rho_n|^{2k}\,d\psi \right]^{r/2k}.$$

where

$$H_i = \int_a^b l_i(x)\, d\psi, \tag{9}$$

$$R_n(f) = \int_a^b \rho_n d\psi. \tag{10}$$

Our method yields results for more general $f(x)$ and for the most general $d\psi(x)$. It further enables us to extend these results to the case of an infinite interval, also to obtain new results in the theory of OP.

1. **Proof of the main result.** Our first objective is to show that

$$\lim_{n\to\infty} \int_a^b \rho_n^2\, d\psi = 0. \tag{11}$$

Note the following properties of MQF (8):

$$H_i = \int_a^b l_i(x) d\psi = \int_a^b l_i^2(x)\, d\psi \qquad (i = 1, 2, \cdots, n), \tag{12}$$

$$R_n(G_{2n-1}) \equiv 0, \tag{13}$$

where $G_s(x) = \sum^s g_i x^i$ generally denotes an arbitrary polynomial of degree $\leqq s$;

$$\lim_{n\to\infty} R_n(f) = 0, \qquad \text{for any } f(x) \left[\text{such that } \int_a^b f(x)\, d\psi \text{ exists}\right]. \tag{14}$$

$$\int_a^b l_i(x) l_j(x)\, d\psi = \int_a^b \frac{\varphi_n(x)}{\varphi_n'(x_i)\, \varphi_n'(x_j)} \frac{\varphi_n(x)}{(x - x_i)(x - x_j)}\, d\psi = 0 \tag{15}$$
$$(i \neq j;\ i, j = 1, 2, \cdots, n).$$

Formulae (12, 13, 15) follow directly from the orthogonality property

$$\int_a^b \varphi_n(x)\, G_{n-1}(x)\, d\psi = 0 \qquad (n = 1, 2, \cdots). \tag{16}$$

We now get from (3), making use of (12, 15):

$$\int_a^b \rho_n^2\, d\psi = \int_a^b f^2(x)\, d\psi - 2\int_a^b f(x) \sum_{i=1}^n l_i(x) f(x_i)\, d\psi + \sum_{i=1}^n f^2(x_i) \int_a^b l_i^2(x)\, d\psi$$

$$\int_a^b \rho_n^2\, d\psi = \int_a^b f^2(x)\, d\psi - 2\int_a^b f(x)[f(x) - L_n(f)]\, d\psi + \sum_{i=1}^n H_i f^2(x_i)$$

$$\int_a^b [f(x) - \rho_n]^2\, d\psi \equiv \int_a^b L_n^2(f)\, d\psi = \sum_{i=1}^n H_i f^2(x_i). \tag{17}$$

By virtue of the convergence property (14),

$$\lim_{n\to\infty} \sum_{i=1}^n H_i f^2(x_i) = \int_a^b f^2(x)\, d\psi. \tag{18}$$

Formula (17) now leads to

$$\lim_{n\to\infty} \int_a^b L_n^2(f)\,d\psi = \int_a^b f^2(x)\,d\psi \tag{19}$$

$$\int_a^b \rho_n^2\,d\psi - 2\int_a^b f(x)\rho_n\,d\psi \to 0, \qquad \text{as } n \to \infty. \tag{20}$$

It is of interest to observe that for a continuous $f(x)$ we may establish the desired formula (11) without resorting to the convergence property (14) of MQF (8). In fact, here to any preassigned $\epsilon > 0$ there corresponds an approximation polynomial (Weierstrass) $P_m(x)$, of proper degree m, such that

$$| f(x) - P_m(x) | \leqq | F(x) | < \epsilon, \qquad a \leqq x \leqq b. \tag{21}$$

We now keep m fixed and take in (3) $n > m$. Since

$$\rho_n(G_{n-1}) = 0, \tag{22}$$

we get

$$\rho_n(f) = F(x) - \sum_{i=1}^{n} l_i(x)F(x_i) \tag{23}$$

and, by the preceding reasoning (see (17)),

$$\int_a^b \rho_n^2\,d\psi - 2\int_a^b \rho_n F(x)\,d\psi + \left[\int_a^b F^2(x)\,d\psi - \sum_{i=1}^{n} H_i F(x_i)\right] = 0. \tag{24}$$

Moreover, by Schwartz's inequality,

$$\int_a^b F(x)\rho_n\,d\psi = \theta_n \sqrt{\left(\int_a^b F^2(x)\,d\psi \cdot \int_a^b \rho_n^2\,d\psi\right)}, \qquad | \theta_n | \leqq 1,$$

and, by virtue of (21),

$$\int_a^b F^2(x)\,d\psi < \epsilon^2\alpha_0, \qquad \alpha_0 = \int_a^b d\psi$$

$$\sum_{i=1}^{n} H_i F^2(x_i) < \epsilon^2 \sum_{i=1}^{n} H_i = \epsilon^2\alpha_0, \qquad (\text{take } f(x) \equiv 1 \text{ in } (8)),$$

and (24) becomes

$$I_n^2 + 2\theta_n' \epsilon \sqrt{(\alpha_0)} I_n + 2\theta_n'' \epsilon^2 \alpha_0 = 0, \qquad I_n^2 = \int_a^b \rho_n^2\,d\psi \tag{25}$$

$$(| \theta_n' |, | \theta_n'' | \leqq 1; \qquad n \geqq N(\epsilon) > m),$$

which, in view of the arbitrariness of ϵ, yields the desired formula (11), for any continuous $f(x)$. Note that (11) implies (20), which in turn, through (17), implies (18, 19).

It is now easy to extend (11) to any $f(x)$. Corresponding to any preassigned

$\epsilon > 0$, we may choose a continuous $\varphi(x)$ and a polynomial $P_m(x)$, of proper degree m, such that

$$\int_a^b [f(x) - \varphi(x)]^2 \, d\psi < \frac{\epsilon}{4}, \qquad \max_{a \leq x \leq b} |\varphi(x) - P_m(x)| < \tfrac{1}{2}\left(\frac{\epsilon}{b-a}\right)^{\frac{1}{2}}.$$

Then

$$\begin{aligned} \int_a^b F^2(x)\, d\psi \equiv \int_a^b [f(x) - P_m(x)]^2 \, d\psi \leqq 2\Bigg\{\int_a^b [f(x) - \varphi(x)]^2 \, d\psi \\ + \int_a^b [\varphi(x) - P_m(x)]^2 \, d\psi\Bigg\} < \epsilon, \end{aligned} \tag{26}$$

and we may now repeat the preceding argument, starting with (24). The only modification necessary is a new estimate of the sum $\sum_{i=1}^n H_i F^2(x_i)$. Here m being fixed, this sum converges, as $n \to \infty$, to $\int_a^b F^2(x)\, d\psi$, so that

$$\sum_{i=1}^n H_i F^2(x_i) = \int_a^b F^2(x)\, d\psi + \epsilon_n, \qquad \lim_{n\to\infty} \epsilon_n = 0; \qquad \epsilon_n < \epsilon \text{ for } n \geqq N(\epsilon) > m,$$

and we get a relation essentially the same as (25):

$$I_n^2 + 2\theta_n' \epsilon I_n + 2\theta_n'' \epsilon = 0 \qquad (|\theta_n'|, |\theta_n''| \leqq 1; n \geqq N(\epsilon) > m),$$

which again yields (11). Finally, using the Minkowsky-Hölder inequality as in footnote 4, we derive our main result:

$$\lim_{n\to\infty} \int_a^b |P_n|^r \, d\psi \equiv \lim_{n\to\infty} \int_a^b |f(x) - L_n(f)|^r \, d\psi = 0, \qquad 0 < r \leqq 2. \tag{27}$$

We have further

$$\left\{\int_a^b |f(x)|^r \, d\psi\right\}^{1/r} \leqq \left\{\int_a^b |f(x) - L_n(f)|^r \, d\psi\right\}^{1/r} + \left\{\int_a^b |L_n(f)|^r \, d\psi\right\}^{1/r}.$$

Hence, by (27),

$$\int_a^b |f(x)|^r \, d\psi \leqq \varliminf_{n\to\infty} \int_a^b |L_n(f)|^r \, d\psi.$$

Similarly, interchanging $f(x)$ and $L_n(f)$,

$$\varlimsup_{n\to\infty} \int_a^b |L_n(f)|^r \, d\psi \leqq \int_a^b |f(x)|^r \, d\psi,$$

whence

$$\lim_{n\to\infty} \int_a^b |L_n(f)|^r \, d\psi = \int_a^b |f(x)|^r \, d\psi, \qquad 0 < r \leqq 2. \tag{27.1}$$

Thus, the sequence $\{L_n(f)\}$ converges to $f(x)$ on (a, b) "strongly," with exponent r.[5] It follows that a subsequence $\{L_{n_p}(f)\}$ of the interpolating poly-

[5] Hobson, *Theory of Functions of a Real Variable*, 2-d ed., 1926, v. 11, pp. 251, 245.

nomials in (4) can be defined which converges to $f(x)$ almost everywhere on (a, b), and that

$$\lim_{n\to\infty} \int_a^b L_n(f)g(x)\,d\psi = \int_a^b f(x)g(x)\,d\psi, \qquad \text{if } \int_a^b g(x)\,d\psi \text{ exists.} \tag{27.2}$$

2. **A more general choice of abscissas.** The following are the salient points in the above analysis:

$$\int_a^b l_i(x)\,d\psi = \int_a^b l_i^2(x)\,d\psi, \qquad i = 1, 2, \cdots, n \tag{12}$$

$$\int_a^b l_i(x)l_j(x)\,d\psi = 0, \qquad i \neq j,\ i, j = 1, 2, \cdots, n \tag{15}$$

$$\lim_{n\to\infty} \sum_{i=1}^n H_i f^2(x_i) = \int_a^b f^2(x)\,d\psi. \tag{18}$$

But (15) is satisfied if (16) is replaced by the less stringent condition

$$\int_a^b \varphi_n(x)G_{n-2}(x)\,d\psi = 0.$$

Furthermore, if a polynomial $\omega_n(x)$, of degree n, satisfies the condition

$$\int_a^b \omega_n(x)G_{n-2}(x)\,d\psi = 0, \tag{28}$$

then necessarily (disregarding constant factors)

$$\omega_n(x) = \varphi_n(x) + C\varphi_{n-1}(x), \qquad C = \text{arbitrary const.},$$

and by a classical argument, all zeros of $\omega_n(x)$ are real and distinct, with one, at most, outside the open interval (a, b). Moreover, (28) *implies* (12). In fact, $l_i^2(x) - l_i(x)$, of degree $2n - 2$, vanishes at $x = x_1, x_2, \cdots, x_n$, so that

$$l_i^2(x) - l_i(x) = \omega_n(x)Q_{n-2}(x),$$

where $Q_{n-2}(x)$ is a polynomial of degree $n - 2$; hence,

$$\int_a^b [l_i^2(x) - l_i(x)]\,d\psi = \int_a^b \omega_n(x)Q_{n-2}(x)\,d\psi^{6} = 0.$$

As to (18), it also holds, the summation being extended over the x_i on (a, b), as was recently proved by the writer.[7]

We thus conclude that *Theorem A holds if in LIF* (3) *the abscissas are the zeros of the polynomials*

$$\omega_n(x) = \varphi_n(x) + C\varphi_{n-1}(x) \qquad (n = 1, 2, \cdots),$$

[6] Cf. L. Fejér, *Mechanische Quadraturen mit positiven Cotesschen Zahlen*, Mathematische Zeitschrift, v. 37, 1933, 287–309; pp. 302–303.

[7] *On Mechanical Quadratures, in particular, with Positive Coefficients*, to appear in the Transactions of the American Mathematical Society.

where the constant C (perhaps depending on n) is such that all zeros of $\omega_n(x)$ belong to (a, b).

Illustration:

$$\omega_n(x) = X_n(x) - X_{n-1}(x), \quad X_n\text{-Legendre polynomial.}$$

Here

$$-1 < x_1 < x_2 < \cdots < x_{n-1} < x_n = 1.^{8}$$

Next, consider the formal expansion

$$(30) \qquad f(x) \sim \sum_{i=0}^{\infty} A_i \varphi_i(x), \qquad A_i = \int_a^b f(x)\varphi_i(x)\, d\psi.$$

If $f(x)$ is continuous in (a, b), then the polynomial

$$(31) \qquad S_{n-1}(x) = \sum_{i=0}^{n-1} A_i \varphi_i(x),$$

formed by the first n terms of the expansion (3), is an interpolation polynomial for $f(x)$, that is,

$$S_{n-1}(\xi_i) = f(\xi_i), \qquad \xi_i \subset (a, b) \qquad (i = 1, 2, \cdots, n).$$

In fact, writing

$$f(x) = S_{n-1}(x) + r_n, \qquad G_{n-1}(x) = \sum_{i=0}^{n-1} \gamma_i \varphi_i(x) \qquad (\gamma_i = \text{const.}),$$

we conclude that

$$(32) \qquad \int_a^b r_n G_{n-1}(x)\, d\psi = 0,$$

hence, the continuous function $r_n = f(x) - S_{n-1}(x)$ vanishes at n (at least) points $\xi_i \subset (a, b)$. (32), for $C_{n-1} \equiv 1$, shows that (6) is satisfied. As to (11), it is also satisfied, for

$$\int_a^b r_n^2\, d\psi = \int_a^b f^2(x)\, d\psi - \sum_{i=0}^{n-1} A_i^2$$

$$(33) \qquad \int_a^b f^2(x) d\psi = \sum_{i=0}^{\infty} A_i^2, \qquad \text{i.e. } \lim_{n\to\infty} \int_a^b r_n^2\, d\psi = 0\text{—Parseval Formula.}$$

Thus, (31) is another example of an interpolation polynomial for which Theorem A holds true. Note that in this case the interpolating abscissas depend on the function $f(x)$.

[8] More generally,

$$\omega_n(x) = \varphi_n(x) - \frac{\varphi_n(b)}{\varphi_{n-1}(b)}\, \varphi_{n-1}(x).$$

3. **Case of an infinite interval.** If (a, b) is infinite, and the abscissas in (3) are again the zeros of $\varphi_n(x; a, b; d\psi)$, then formulae (12, 15) hold true. However, (18) is not necessarily valid for an infinite interval, nor does there always exist a polynomial $P_m(x)$, of proper degree m, such that

$$\int_a^b [f(x) - P_m(x)]^2 \, d\psi < \epsilon \qquad (\epsilon \text{ positive, arbitrarily small}). \tag{26}$$

Regarding (26), observe that its validity is equivalent to that of Parseval Formula (33). In the first place, (33) may be rewritten as

$$\lim_{n\to\infty} \int_a^b [f(x) - S_{n-1}(x)]^2 \, d\psi = 0; \tag{34}$$

in the second place, it is known from the theory of OP that

$$\int_a^b [f(x) - S_{n-1}(x)]^2 d\psi \leqq \int_a^b [f(x) - G_{n-1}(x)]^2 \, d\psi.$$

As to (18), which states the convergence of the corresponding MQF (8), it is implied, for some kinds of $\psi(x)$, by Parseval Formula (33).[9] Hence, the conclusion: Theorem A holds in all cases (assuming the existence of $\int_a^b f^2(x)\, d\psi$), provided, the Parseval Formula holds for the OP employed, and the MQF converges. [Theorem A then evidently holds also for thepolynomial (31).]

In particular, Theorem A is valid if we use in (3) for abscissas the zeros of Laguerre or Hermite polynomials.

4. **Further study of $l_i(x)$, H_i.** In this section (a, b) may be finite or infinite. The orthogonality property (15) of the $l_i(x)$ leads to many interesting properties of the basic polynomials and of the coefficients H_i in MQF (8).

In view of the linear independence of the $l_i(x)$[10] and by virtue of (15, 12, 22), our LIF (4), for a polynomial of degree $\leqq n - 1$ leads to

$$G_{n-1}(x_i) = \int_a^b G_{n-1}(x) \frac{l_i(x)}{H_i} \, d\psi, \qquad i = 1, 2, \cdots, n. \tag{35}$$

Furthermore,

$$G_n(x) = \frac{g_n \varphi_n(x)}{a_n} + G_{n-1}(x), \qquad G_{n-1}(x_i) = G_n(x_i), \quad i = 1, 2, \cdots, n.$$

Hence, by (35, 16),

---

[9] Jacques Chokhate (J. Shohat), *Sur la convergence des quadratures méchaniques cans un intervalle infini. Applications au problème des moments, au calcul des probabilités.* Comptes Rendus, v. 186, 1928, 344–346.

[10] In fact, a linear relation $\sum_{i=1}^n C_i l_i(x) \equiv 0$ yields, for $x = x_j : C_j = 0, j = 1, 2, \cdots, r$.

$$(36)\qquad H_i = \int_a^b \frac{G_n(x)}{G_n(x_i)}\, l_i(x)\, d\psi \qquad (G_n(x_i) \neq 0), \quad i = 1, 2, \cdots, n.$$

$$(37)\quad G_n(x) = \sum_{i=1}^{n} l_i(x) \int_a^b G_n(x)\, \frac{l_i(x)}{H_i}\, d\psi + \rho_n(G_n), \qquad \rho_n(G_n) = \frac{g_n}{a_n}\, \varphi_n(x),$$

and we thus get the following new interpretation of LIF (3).

THEOREM B. *LIF, using for abscissas the zeros of any OP sequence* $\varphi_n(x; a, b; d\psi)$, *when applied to a polynomial* $G_n(x)$, *of degree* $\leqq n$, *is identical with the expansion of* $G_n(x)$ *according to the limited orthonormal sequence of the basic polynomials* $\{l_i(x)/H_i^{\frac{1}{2}}\}$ $(i = 1, 2, \cdots, n)$, *all of degree* $n - 1$.

Formula (37) yields at once

$$(38)\qquad \int_a^b G_n^2(x)\, d\psi = \sum_{i=1}^{n} H_i G_n^2(x_i) + \frac{g_n^2}{a_n^2},$$

so that the following relation holds

$$\int_a^b \rho_n(G_n^2)\, d\psi = \int_a^b \rho_n^2(G_n)\, d\psi.$$

We proceed to draw various conclusions from the formulae just established.

First, (38) shows that

$$\rho_n(G_n^2) \geqq 0, \text{ equality if and only if } g_n = 0,$$

that is, MQF (8) cannot be exact for a polynomial of degree n of the form $G_n^2(x)$; more precisely,

$$\int_a^b G_n^2(x)\, d\psi \geqq \sum_{i=1}^{n} H_i G_n^2(x_i), \quad \text{equality if and only if } g_n = 0.$$

We further conclude that

$$(39)\qquad H_i = \min. \int_a^b \frac{G_n^2(x)}{G_n^2(x_i)}\, d\psi, \qquad i = 1, 2, \cdots, n,$$

minimum attained if and only if

$$g_n = 0, \qquad G_n(x_j) = 0, \qquad 1 \leqq j \leqq n, \qquad j \neq i,$$

that is for $G_n(x) = \text{Const.} \times l_i(x)$ only.

Secondly, various results may be obtained by specifying in various manners $G_n(x)$ in (36). Thus, take in (36) $G_n(x) \equiv \varphi_{n-1}(x)$, and we get

$$(40)\qquad H_i = \frac{a_n}{a_{n-1}\varphi_{n-1}(x_i)\varphi_n'(x_i)} = \frac{1}{K_n(x_i)} = \frac{1}{K_{n-1}(x_i)}, \quad i = 1, 2, \cdots, n,$$

where, by Darboux formula,

$$K_n(x) \equiv \sum_{i=0}^{n} \varphi_i^2(x) = \frac{a_n}{a_{n+1}}\, [\varphi_{n+1}'(x)\varphi_n(x) - \varphi_n'(x)\varphi_{n+1}(x)].$$

Since all $H_i > 0$, it follows from (40) that

$$\varphi_{n-1}(x_i)\varphi_n'(x_i) > 0, \qquad i = 1, 2, \cdots, n,$$

so that the zeros of $\varphi_{n-1}(x)$ separate those of $\varphi_n(x)$.

Assuming $a \geqq 0$, take in (36) $G_n(x) \equiv x\varphi_{n-1}(x; d\psi_1)$ where generally

$$d\psi_k(x) \equiv x^k \, d\psi(x). \qquad (k > 0). \tag{41}$$

We obtain

$$H_i = a_n(d\psi)/a_{n-1}(d\psi_1)x_i\varphi_{n-1}(x_i\,; d\psi_1)\varphi_n'(x_i\,; d\psi), \qquad i = 1, 2, \cdots, n. \tag{42}$$

Since here not only H_i, but also x_i is > 0 for $i = 1, 2, \cdots, n$, we conclude, as above, that the zeros of $\varphi_{n-1}(x; d\psi_1)$ separate those of $\varphi_n(x; d\psi)$. Moreover, a comparison of (42) and (40) shows that

$$a_{n-1}(d\psi_1)x\varphi_{n-1}(x; d\psi_1) - a_{n-1}(d\psi)\varphi_{n-1}(x; d\psi) = 0 \text{ at } x = x_1\,, x_2\,, \cdots, x_n\,,$$

whence,

$$x\varphi_{n-1}(x; d\psi_1) = \frac{a_{n-1}(d\psi)}{a_{n-1}(d\psi_1)}\,\varphi_{n-1}(x; d\psi) + \frac{a_{n-1}(d\psi_1)}{a_n(d\psi)}\,\varphi_n(x; d\psi).$$

Returning to the case of any (a, b), we obtain another expression for H_i, by taking in (36) $G_n(x) = x^2\varphi_{n-2}(x; d\psi_2)$ and making use of the relation

$$\int_a^b G_{n-1}(x)\varphi_n(x)\,d\psi = (g_{n+1}S_{n+1}(d\psi) + g_n)/a_n(d\psi):$$

$$H_i = a_n(d\psi)[S_{n-1}(d\psi_2) - S_n(d\psi) + x_i]/x_i^2\varphi_{n-2}(x_i\,; d\psi_2)\varphi_n'(x_i\,; d\psi)a_{n-2}(d\psi_2), \tag{43}$$
$$i = 1, 2, \cdots, n \qquad (x_i \neq 0).$$

For symetric OP $[(a, b) \equiv (-h, h),\ \psi(-x) \equiv -\psi(x),\ S_n(d\psi) = 0]$ (43) becomes

$$H_i = a_n(d\psi)/x_i\varphi_{n-2}(x_i\,; d\psi_2)\varphi_n'(x_i\,; d\psi)a_{n-2}(d\psi_2), \qquad i = 1, 2, \cdots, n \quad (x_i \neq 0). \tag{43.1}$$

We see that the orthogonality property (15) of the $l_i(x)$ yields very readily many properties of OP some of which are new, some have been derived in a more complicated manner. (39) represents an improvement over an earlier result where in place of n we find $n - 1$. This improvement is important for it enabled us to choose in (36) as $G_n(x)$ a polynomial *precisely of degree n*. Observe that the same formula (36) yields bounds for x_i, if such are known for H_i, and vice versa.[11]

[11] Thus, for example,

$$H_i < \frac{\alpha_{2s}}{x_i^{2s}} \quad \left(i = 1, 2, \cdots, n;\ 0 \leqq s \leqq n;\ \alpha_k = \int_a^b x^k\,d\psi\right).$$

Cf. J. Shohat and C. Winston, *On Mechanical Quadratures*, Rendiconti del Circolo Matematico di Palermo, v. 58, 1934, 1–137, where the above formula was made use of, with $s \leqq n - 1$.

5. **Application to classical OP.** Apply the above results to the polynomials (L) of Laguerre and (J) of Jacobi, where

$$(L) \quad (a, b) \equiv (0, \infty), \quad d\psi = x^{\alpha-1}e^{-x}\,dx, \quad \alpha > 0; \quad S_n = n(n + \alpha - 1),$$
$$a_n = [\Gamma(n+1)\Gamma(n+\alpha)]^{-\frac{1}{2}}$$

$$(J) \quad (a, b) = (0, 1), \quad d\psi = x^{\alpha-1}(1-x)^{\beta-1}, \quad \alpha, \beta > 0, \quad S_n = \frac{n(n+\alpha-1)}{2n+\alpha+\beta-2}.$$

(i) *Polynomials of Laguerre.* Here

$$(44) \qquad H_i = \frac{1}{x_i^2\sqrt{(n(n-1))}}\,\frac{x_i - \alpha}{\varphi_{n-2}(x_i\,;\,d\psi_2)\varphi_n'(x_i\,;\,d\psi)} \quad (i = 1, 2, \cdots, n;\ n \geqq 2).$$

It follows that

$$(45) \qquad x_i \neq \alpha,\ x_1 < \alpha \qquad (i = 1, 2, \cdots, n;\ n \geqq 2)$$

$$(46) \qquad \varphi_{n-2}(x_i\,;\,d\psi_2)\varphi_n'(x_i\,;\,d\psi) \lessgtr 0, \qquad \text{according as } x_i \lessgtr \alpha,$$

that is, if j_n zeros of $\varphi_n(x;\,d\psi)$ do not exceed α,

$$x_1 < x_2 < \cdots < x_{j_n} < \alpha < x_{j_n+1} < \cdots < x_n,$$

then $j_n - 1$ zeros of $\varphi_{n-2}(x;\,d\psi_2)$ separate the first j_n zeros of $\varphi_n(x;\,d\psi)$, the remaining $n - j_n - 1$ zeros of $\varphi_{n-2}(x;\,d\psi_2)$ separate the last $n - j_n$ zeros of $\varphi_n(x;\,d\psi)$, while the interval (x_{j_n}, x_{j_n+1}) contains no zeros of $\varphi_{n-2}(x;\,d\psi_2)$.

Apply the above considerations to

$$\varphi_n(x;\,d\psi), \qquad \varphi_{n-2}(x;\,d\psi_2), \qquad \varphi_{n-4}(x;\,d\psi_4), \cdots$$

and notice that the $\varphi_1(x;\,d\psi_k)$ vanishes at $x = \alpha + k$. We conclude that

$$(47) \qquad j_n \leqq \left[\frac{n}{2}\right].$$

(ii) *Jacobi Polynomials.* Here

$$(48) \qquad H_i = \sqrt{\frac{(\alpha+\beta+2n-1)(\alpha+\beta+2n-2)^2(\alpha+\beta+2n-1)}{n(n-1)(n+\beta-1)(n+\beta-2)}} \cdot \frac{x_i - \dfrac{\alpha}{2n+\alpha+\beta-2}}{x_i^2\varphi_{n-2}(x_i\,;\,d\psi_2)\varphi_n'(x_i\,;\,d\psi)} \quad (i = 1, 2, \cdots, n;\ n \geqq 2),$$

and the above considerations holds, with α replaced by $\alpha/(2n + \alpha + \beta - 2)$. The inequality

$$(49) \qquad x_{1,n} \equiv x_1 < \frac{\alpha}{2n+\alpha+\beta-2} \qquad (n \geqq 2)$$

is especially interesting, in view of the known property: $\lim_{n\to\infty} x_{1,n} = 0$.

Notice that

$$(L)\ x_{1,1} = \alpha, \qquad (J)\ x_{1,1} = \frac{\alpha}{\alpha + \beta},$$

so that

$$(L, J) \qquad j_n = \left[\frac{n}{2}\right] \qquad \text{for } n = 3.$$

(Added in proof.) Professor J. D. Tamarkin made the following observation. If we consider $L_m(f)$ as a linear transformation on the space C of functions continuous on $(-1, 1)$ to the space L_2, it follows from form. (39), p. 154, of the above paper by Erdös and Turán that the modulus of this transformation is $\geq S(m)$. Since this quantity is assumed not to be bounded, it follows immediately from a well known theorem of Banach (Theórie des operations linéaires, p. 80, Th. 5) that the sequence $\{L_m(f)\}$ cannot be bounded (in L_2), and thus there exists a continuous function $f(x)$ for which $\limsup_m \int_{-1}^{1} L_m^2(f)\,dx = \infty$.

THE UNIVERSITY OF PENNSYLVANIA.

ANNALS OF MATHEMATICS
Vol. 38, No. 4, October, 1937

A REPRESENTATION OF ALL ANALYTIC FUNCTIONS IN TERMS OF FUNCTIONS WITH POSITIVE REAL PART

BY M. S. ROBERTSON

(Received April 15, 1937)

Introduction. In a recent paper[1] the author discussed a class of analytic functions

$$f(z) = z + \sum_{2}^{\infty} a_n z^n$$

regular for $|z| < 1$ having the property that for all values of r less than and near to one, $f(z)$ mapped $|z| = r$ into a contour which cut across the real axis in exactly two points. Such a function was called *star-like in the direction of the real axis*. It was shown that $f(z)$ could be represented in the form

$$f(z) = h_\nu(ze^{-i\mu})[\cos\mu + i\sin\mu \cdot F(z)]$$

where μ and ν are certain constants dependent upon $f(z)$ with $\sin\mu \geqq 0$. Here $F(z)$ is analytic for $|z| < 1$, has a positive real part for $|z| < 1$, and $F(0) = 1$. The function

$$h_\nu(z) \equiv z(1 - 2z\cos\nu + z^2)^{-1}.$$

From this representation it was deduced that $|a_n| \leqq n^2$ and that the equality sign was attained by essentially only one function of this class.

It is now our purpose in this paper to generalize these results to analytic functions having the first several coefficients of the power series all zero. We may remark that every function which is analytic about the origin can be so normalized that it is holomorphic within the unit circle, vanishes at the origin, and has its first non-vanishing coefficient unity. Hence we may assume that our function has the expansion

$$f_k(z) = z^k + \sum_{k+1}^{\infty} a_n z^n, \qquad k \geqq 1. \tag{1.1}$$

Now for small values of $|z|$ this function behaves very much like the first term z^k which maps a circle into another circle the circumference of which is traced out k times. Thus for sufficiently small values of r the circle $|z| = r$ is mapped by the function $f_k(z)$ into a contour C_r which consists of k loops about the origin, each loop cutting the real axis in two points, one point on either side of the origin. In all C_r cuts the real axis in exactly $2k$ points. By a magnification

[1] See M. S. Robertson, "Analytic functions star-like in one direction," *American Journal of Mathematics*, Vol. LVIII, No. 3, (1936), pp. 465–472.

of the variable we can easily arrange that $f_k(z)$ has this property within the whole unit circle. Hence we shall assume this to be the case for $f_k(z)$ of (1.1) and denote by (S_k) the class of all such functions.

For these functions we show that $f_k(z)$ can always be represented in the following form:

$$(1.2) \qquad f_k(z) = \frac{z^k}{\prod_{s=1}^{2k} (1 - ze^{-i\theta_s})} \cdot [1 + ie^{-i\sigma_k} \cdot \sin \sigma_k \cdot \{F(z) - 1\}]$$

where θ_s are $2k$ constants dependent upon $f_k(z)$ with

$$(-1)^{k-1} \sin \sigma_k \geqq 0, \qquad \sigma_k = \tfrac{1}{2} \sum_1^{2k} \theta_s,$$

and where $F(z)$ is regular for $|z| < 1$, $\Re F(z) > 0$ for $|z| < 1$, and $F(0) = 1$.

From this representation for $f_k(z)$ we deduce that

$$(1.3) \qquad |a_n| \leqq \frac{n}{k} \cdot \frac{(n + k - 1)!}{(2k - 1)!\,(n - k)!} = O(n^{2k})$$

where the equality sign is attained by essentially only one function

$$\frac{z^k + \epsilon z^{k+1}}{(1 - \epsilon z)^{2k+1}}, \qquad \epsilon = \pm e^{i\pi/2k}.$$

In particular, if $k = 1$, we have $|a_n| \leqq n^2$.

From (1.2) it will also follow that

$$(1.4) \qquad \lim_{r \to 1} f_k(re^{i\theta}) = f_k(\theta)$$

exists as a finite limit almost everywhere.

In section five it is shown that for the functions $f_k(z)$ which are also real on the real axis, formula (1.2) simplifies to the much neater form:

$$(1.5) \qquad f_k(z) = \frac{z^k F(z)}{(1 - z^2) \cdot \prod_{j=2}^{k} (1 - 2z \cos \theta_j + z^2)}.$$

In this case there are only $(k - 1)$ real parameters θ_j dependent upon $f_k(z)$ instead of the usual $2k$ parameters.

If

$$(1.6) \qquad w = g_k(z) = z + \sum_2^{\infty} C_n^{(k)} z_n$$

is regular for $|z| < 1$, and in the unit circle is also star-like in the direction of the $2k$ rays

$$\arg w = \frac{p}{k}\pi, \qquad p = 0, 1, 2, \cdots, 2k - 1$$

then it is shown that $[g_k(z)]^k$ belongs to the class (S_k) considered above. From this fact it is found that the coefficients $C_n^{(k)}$ satisfy the inequalities

$$| C_n^{(k)} | \leqq d_n^{(k)} < en^{1+(1/k)}$$

where $d_n^{(k)}$ is the coefficient of z^n in the power series for the function $\dfrac{z(1+z)^{1/k}}{(1-z)^{2+(1/k)}}$. In particular, if $g_k(z)$ is star-like in every direction we have for its coefficients $C_n^{(\infty)}$ the inequality $| C_n^{(\infty)} | \leqq n$, which is well-known.

2. **The imaginary part of $f_k(z)$.** We shall now give an analytical characterization of this geometric property of the functions of the type mentioned above. Let (S_k) denote the class of functions $f_k(z)$ which have one or the other of the following sequence of properties:

Either (A): 1. $f_k(z) = z^k + \sum_{k+1}^{\infty} a_n z^n$ is regular for $| z | < 1$.

2. There exists a $\delta = \delta(f_k)$ so that for every r in the open interval $1 - \delta < r < 1$ $f_k(z)$ maps $| z | = r$ into a contour C_r which cuts across the real axis in $2k$, and not more than $2k$ points.

or (B): 1. $f_k(z) = z^k + \sum_{k+1}^{\infty} a_n z^n$ is regular for $| z | \leqq 1$ except for at most a finite number of poles on $| z | = 1$.

2. $f_k(z)$ maps $| z | = 1$ into a contour which cuts across the rea axis in exactly $2k$ points.

Corresponding to the points where C_r cuts the real axis, there will be $2k$ points $z_j = re^{i\bar{\theta}_j(r;f)}$, $j = 1, 2, \cdots 2k$, at which the imaginary part of $f_k(z)$, or $\Im f_k(z)$ is zero. We may define $\bar{\theta}_j(r) \equiv \bar{\theta}_j(r;f)$ so that

$$(2.1) \qquad \begin{aligned} &0 \leqq \bar{\theta}_1(r) \leqq 2\pi, \quad 0 \leqq \bar{\theta}_1(r) < \bar{\theta}_2(r) < \cdots < \bar{\theta}_{2k}(r) \\ &0 < \bar{\theta}_i(r) - \bar{\theta}_j(r) < 2\pi, \ i > j; \ i, j = 1, 2, \cdots 2k. \end{aligned}$$

If we set

$$f_k(re^{i\theta}) = u(r, \theta) + i\, v(r, \theta)$$

then, with a proper choice of $\bar{\theta}_1(r)$,

$$(2.2) \qquad \begin{aligned} &v(r, \theta) \geqq 0 \text{ when } \bar{\theta}_{2s-1}(r) < \theta < \bar{\theta}_{2s}(r), \qquad s = 1, 2, \cdots k, \\ &v(r, \theta) \leqq 0 \text{ when } \bar{\theta}_{2s}(r) < \theta < \bar{\theta}_{2s+1}(r), \qquad s = 1, 2, \cdots k, \end{aligned}$$

where $\bar{\theta}_{2k+1}(r) \equiv \bar{\theta}_1(r) + 2\pi$.

For brevity we shall denote here α_s, $\bar{\mu}_s$, $\bar{\nu}_s$, ϕ_s by

$$\alpha_s = \pi - \left(\frac{\bar{\theta}_{2k-s+2} - \bar{\theta}_{s+1}}{2}\right), \qquad s = 2, 3, \cdots, k.$$

$$\bar{\mu}_s = \frac{\bar{\theta}_{2k-s+2} + \bar{\theta}_{s+1}}{2}, \qquad s = 2, 3, \cdots, k,$$

$$\bar{\nu}_s = \frac{\bar{\theta}_{2k-s+2} - \bar{\theta}_{s+1}}{2} = \pi - \alpha_s, \qquad s = 2, 3, \cdots, k, \tag{2.3}$$

$$\phi_1 = \frac{\pi}{2} - \bar{\mu}_1, \quad \bar{\mu}_1 = \frac{\bar{\theta}_1 + \bar{\theta}_2}{2}, \qquad \bar{\nu}_1 = \frac{\bar{\theta}_2 - \bar{\theta}_1}{2},$$

$$\phi_s = \frac{3\pi}{2} - \phi_1 - \phi_2 - \cdots - \phi_{s-1} - \left(\frac{\bar{\theta}_{2k-s+2} + \bar{\theta}_{s+1}}{2}\right).$$

The functions defined above are clearly functions of r, while s takes on the values $2, 3, \cdots, k$. With this notation we are now in a position to obtain the formula (1.2) for $f_k(z)$.

3. A representation for the functions of class (S_k). Let r be any number in the range $1 - \delta < r < 1$. In order to bring the point whose argument is $\bar{\mu}_1(r) = \frac{1}{2}(\bar{\theta}_1(r) + \bar{\theta}_2(r))$ to the positive imaginary axis we operate upon $f(z)$ with the rotational function $e^{-i\phi_1(r)}$ where $\phi_1(r) = \frac{1}{2}\pi - \bar{\mu}_1(r)$, forming

$$H_1(z) \equiv \frac{f_k(rze^{-i\phi_1})}{h_{\bar{\nu}_1}(-iz)} = ir^k e^{-ki\phi_1} \cdot z^{k-1} + \cdots \tag{3.1}$$

where $h_\nu(z) \equiv z(1 - 2z\cos\nu + z^2)^{-1}$. $H_1(z)$ is regular for $|z| \leqq 1$. For $z = e^{i\theta}$ we have

$$\Im H_1(z) = 2\{\Im f_k(rze^{-i\phi_1})\} \cdot \{\sin\theta - \sin(\tfrac{1}{2}\pi - \bar{\nu}_1)\}. \tag{3.2}$$

An examination of the intervals in which these two factors have the same sign will reveal that $H_1(z)$ (save for a rotation and magnification to make the leading coefficient of the power series unity) is of class (S_{k-1}). Introducing now a second rotational function $e^{-i\phi_2(r)}$ where

$$\phi_2(r) = \frac{\pi}{2} - \left(\frac{\bar{\theta}_3 + \bar{\theta}_{2k}}{2} + \phi_1 - \pi\right) = \frac{3\pi}{2} - \phi_1 - \left(\frac{\bar{\theta}_3 + \bar{\theta}_{2k}}{2}\right)$$

we repeat the process, forming the function

$$H_2(z) \equiv \frac{H_1(ze^{-i\phi_2})}{h_{\alpha_2}(-iz)} = (i)^2 r^k e^{-(k\phi_1 + \overline{k-1}\phi_2)i} \cdot z^{k-2} + \cdots. \tag{3.3}$$

Again (save for the first coefficient), $H_2(z)$ is of class (S_{k-2}). This operation, which we have now performed twice, is performed k times in all. At the j^{th} stage we have

$$H_j(z) = \frac{H_{j-1}(ze^{-i\phi_j})}{h_{\alpha_j}(-iz)} = (i)^j r^k e^{-i\sum_{s=1}^{j}(k+1-s)\phi_s} \cdot z^{k-j} + \cdots. \tag{3.4}$$

$H_j(z)$ (save for the first coefficient) is of class (S_{k-j}) and there are now $(k - j)$ intervals on the unit circle for which $\Im H_j(e^{i\theta}) \geqq 0$ and the same number for which $\Im H_j(e^{i\theta}) \leqq 0$. Hence at the k^{th} stage we have

$$H_k(z) = \frac{H_{k-1}(ze^{-i\phi_k})}{h_{\alpha_k}(-iz)} = (i)^k \cdot r^k \cdot e^{-i\sum_{s=1}^{k}(k+1-s)\phi_s} + b_1 z + \cdots \tag{3.5}$$

with

$\Im\, H_k(z) \geqq 0$ for $|z| = 1$, and also for $|z| < 1$ since the minimum value occurs on the boundary.

Since
$$\Im\, H_k(0) = \Im[(ir)^k e^{-i\sum_{s=1}^{k}(k+1-s)\phi_s}] \geqq 0$$
and since from (2.3) we have
$$\sum_{s=1}^{k}(k+1-s)\phi_s = (3k-2)\frac{\pi}{2} - \sum_{1}^{k}\bar{\mu}_s$$
we obtain
$$(-1)^{k-1}\cdot\sin\left(\sum_{1}^{k}\bar{\mu}_s\right) \geqq 0. \tag{3.6}$$
Also we have
$$\begin{aligned} f_k(rz) &= H_1(ze^{i\phi_1})\cdot h_{\bar{\nu}_1}(-ie^{i\phi_1}\cdot z) \\ &= H_2\{ze^{i(\phi_1+\phi_2)}\}\cdot h_{\alpha_2}\{-ize^{i(\phi_1+\phi_2)}\}\cdot h_{\bar{\nu}_1}(-ize^{i\phi_1}) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &= H_k(ze^{i\sum_1^k\phi_s})\cdot h_{\bar{\nu}_1}(-ize^{i\phi_1})\cdot\prod_{j=2}^{k} h_{\alpha_j}\{-ize^{i\sum_1^j\phi_s}\} \\ &= H_k(-ize^{-i\bar{\mu}_k})\cdot h_{\bar{\nu}_1}(ze^{-i\bar{\mu}_1})\cdot\prod_{j=2}^{k} h_{\alpha_j}(-ze^{-i\bar{\mu}_j}) \\ &= (-1)^{k-1}\cdot H_k(-ize^{-i\bar{\mu}_k})\cdot\prod_{j=1}^{k} h_{\bar{\nu}_j}(ze^{-i\bar{\mu}_j}). \end{aligned} \tag{3.7}$$
Let $\{r_n\}$ be a sequence of values of $r < 1$ and tending to unity in such a way that the following limits exist:
$$\begin{aligned} &\lim_{r_n\to 1}\bar{\theta}_j(r_n) = \theta_j, && j = 1, 2, \cdots, 2k. \\ &\lim_{r_n\to 1}\bar{\mu}_s(r_n) = \mu_s, && s = 1, 2, \cdots, k \\ &\lim_{r_n\to 1}\bar{\nu}_s(r_n) = \nu_s, && s = 1, 2, \cdots, k \\ &\lim_{r_n\to 1}H_k(z) \equiv H(z), && |z| < 1. \end{aligned} \tag{3.8}$$
Hence in proceeding to the limit in (3.7) we have
$$f_k(z) = (-1)^{k-1}H(-ize^{-i\mu_k})\cdot\prod_{j=1}^{k} h_{\nu_j}(ze^{-i\mu_j}) \tag{3.9}$$
where $H(z)$ is regular for $|z| < 1$, and $\Im\, H(z) \geqq 0$ for $|z| < 1$. Also proceeding to the limit in (3.6) we have
$$(-1)^{k-1}\sin\left(\sum_{1}^{k}\mu_s\right) \geqq 0. \tag{3.10}$$
Case 1. Suppose $\sin(\sum_1^k \mu_s) \neq 0$. We define

(3.11) $$F(z) \equiv \frac{(-1)^{k-1}\, i \cos\left(\sum_1^k \mu_s\right) - i\, H(-ize^{-i\mu_k})}{(-1)^{k-1} \sin\left(\sum_1^k \mu_s\right)}.$$

Then $F(z)$ is regular for $|z| < 1$, $\Re F(z) > 0$ for $|z| < 1$, $F(0) = 1$. We can now write $f(z)$ in the final form

(3.14) $$\begin{aligned} f_k(z) &= \prod_{j=1}^k h_{\nu_j}(ze^{-i\mu_j}) \cdot \left[\cos\left(\sum_1^k \mu_s\right) + i \sin\left(\sum_1^k \mu_s\right) \cdot F(z)\right] \\ &= \frac{z^k}{\prod_{s=1}^{2k} (1 - ze^{-i\theta_s})} [1 + ie^{-i\sigma_k} \cdot \sin \sigma_k \cdot \{F(z) - 1\}] \end{aligned}$$

where $\sigma_k = \frac{1}{2} \sum_1^{2k} \theta_j$.

Case 2. Suppose $\sin\left(\sum_1^k \mu_s\right) = 0$. Then we have $(-1)^{k-1} H(z) \equiv \pm 1$ since $\Im H(0) = 0$. Hence from a comparison of (3.9) and (3.14) we see that the formula (3.14) still holds good for this second case if we define $F(z) \equiv 1$.

We may let

(3.15) $$F(z) = 1 + \sum_1^\infty b_n z^n.$$

Then since $\Re F(z) > 0$ for $|z| < 1$, it is well known[2] that

(3.16) $$|b_n| \leqq 2 \text{ for all } n.$$

Hence it follows that

(3.17) $$|F(re^{i\theta}) - 1| \leqq 2 \sum_1^\infty r^n = \frac{2r}{1-r}, \qquad r < 1.$$

It follows from (3.14) that

(3.18) $$|f_k(re^{i\theta})| \leqq \frac{r^k}{(1-r)^{2k}} \cdot \left[1 + \frac{2r}{1-r}\right] = \frac{r^k(1+r)}{(1-r)^{2k+1}}.$$

Since for analytic functions with positive real part for $|z| < 1$

$$\lim_{r \to 1} F(re^{i\theta}) = F(\theta)$$

exists as a finite limit almost everywhere, from (3.14) it follows that

(3.19) $$\lim_{r \to 1} f_k(re^{i\theta}) = f_k(\theta)$$

exists and is finite almost everywhere.

[2] See C. Carathéodory, *Mathematische Annalen*, Bd. 64, (1907), S. 95–115.

We have seen that every function of class (S_k) can be represented in the form (3.14). Conversely, if $F(z)$ is regular for $|z| \leqq 1$ save for a finite number of poles on $|z| = 1$ and $\Re F(z) \geqq 0$ for $|z| \leqq 1$, and if μ_s, ν_s are any constants subject to the condition (3.10) then $f_k(z)$ obtained by formula (3.14) is a member of class (S_k) and maps the unit circle into a contour cutting the real axis in $2k$ points. If, however, $F(z)$ is not defined on the unit circle we may form $f_k(z)$ as in (3.14) and also the function $f_k^n(z)$ obtained from (3.14) by replacing $F(z)$ by $F(r_n z)$ where $\{r_n\}$ is any sequence of values of $r < 1$ tending to unity. Then each $f_k^n(z)$ is a member of (S_k) while

$$f_k(z) \equiv \lim_{n \to \infty} f_k^n(z)$$

uniformly in any domain completely interior to the unit circle.

4. The coefficients of the power series. Assuming that

$$F(z) = 1 + \sum_1^\infty b_n z^n, \qquad \Re F(z) \geqq 0 \qquad \text{for } |z| < 1, \tag{4.1}$$

we have[3] $|b_n| \leqq 2$ for all n, equality for $n = 1$ holding only for $\dfrac{1 + e^{i\alpha} \cdot z}{1 - e^{i\alpha} \cdot z}$, α real. Hence if we expand each function $h_{\nu_j}(ze^{-i\mu_j})$ in (3.14) in a power series we can identify the coefficients of like powers of z on each side of the equation (3.14). If then we replace each coefficient by its largest absolute value we have

$$(z^k + z^{k+1})(1 - z)^{-2k-1} = z^k + \sum_{k+1}^\infty c_n z^n$$

as a majorant for $f_k(z)$. Consequently we have for $n > k$

$$|a_n| \leqq c_n = \frac{n}{k} \cdot \frac{(n + k - 1)!}{(2k - 1)!\,(n - k)!} = O(n^{2k}). \tag{4.2}$$

In particular if $k = 1$ we have $|a_n| \leqq n^2$. The inequality (4.2) cannot be improved upon as the equality sign is attained by a member of class (S_k), namely,

$$\frac{z^k(1 + \epsilon z)}{(1 - \epsilon z)^{2k+1}}, \qquad \epsilon = \pm e^{i\pi/2k}. \tag{4.3}$$

Moreover, this is the only function of class (S_k) for which the equality sign holds. For we must have all the μ_j equal to one another, $(-1)^{k-1} \sin (\sum_1^k \mu_j) = \frac{1}{2}\pi$, ν_j either all 0 or all equal to π. Whence $\mu_j = e^{i\pi(1-(1/2k))}$ for all j. Hence $f_k(z)$ must have the form

$$f_k(z) = \frac{z^k F(z)}{(1 - \epsilon z)^{2k}}, \qquad \epsilon = \pm e^{i\pi/2k}$$

[3] See W. Rogosinski, "Über Bildschranken bei Potenzreihen und ihren Abschnitten," *Math. Zeit.*, 17 Bd., (1923), s. 265. See also G. Julia: *Principes géométriques d'analyse*, Paris, (1930), p. 107.

and again for equality in (4.2) we must have $|b_1| = 2$ and $F(z) = (1 + \epsilon z)/(1 - \epsilon z)$. Hence equality in (4.2) for any fixed n occurs for essentially the one function (4.3).

5. **Typically-real functions of class k.** The case where $f_k(z)$ is also real on the real axis is of some interest in itself. If $k = 1$ and if $f_1(z)$ is real on the real axis, $f_1(z)$ is then typically-real for $|z| < 1$, following the definition due to W. Rogosinski.[4] In this case since $\theta_1(r) \equiv 0$, $\theta_2(r) \equiv \pi$, $\mu_1 = \nu_1 = \frac{1}{2}\pi$ formula (3.14) simplifies to the form

$$f_1(z) = \frac{zF(z)}{1 - z^2} \tag{5.1}$$

where $F(z)$ is real on the real axis and $\Re F(z) > 0$ for $|z| < 1$. Formula (5.1) was first obtained by W. Rogosinski.[5] As he has shown, the converse theorem is also true: given $F(z)$ regular for $|z| < 1$, $F(0) = 0$, real on the real axis with $\Re F(z) > 0$ for $|z| < 1$, then the function $f_1(z)$ formed as in (5.1) is typically-real with respect to the unit circle. We now generalize this idea to functions of class k.

DEFINITION. If $f(z)$ is a member of class (S_k) and is also real on the real axis, we shall call $f(z)$ a *typically-real function of class k with respect to the unit circle.*

Let $f_k(z)$ be typically-real of class k. Then

$$\theta_1 = 0, \quad \theta_{k+1} = \pi, \quad \theta_{2k-s+2} = 2\pi - \theta_s, \qquad s = 2, 3, \cdots, k.$$

$$\tag{5.2} \begin{aligned} \mu_1 &= \frac{\theta_2}{2}, \quad \mu_s = \frac{\theta_{2k-s+2} + \theta_{s+1}}{2} = \pi + \frac{\theta_{s+1} - \theta_s}{2}, \qquad s = 2, 3, \cdots, k. \\ \nu_1 &= \frac{\theta_2}{2}, \quad \nu_s = \frac{\theta_{2k-s+2} - \theta_{s+1}}{2} = \pi - \frac{\theta_{s+1} + \theta_s}{2}, \qquad s = 2, 3, \cdots, k. \end{aligned}$$

$$\sum_1^k \mu_s = \frac{\theta_2}{2} + \sum_2^k \left(\pi + \frac{\theta_{s+1} - \theta_s}{2}\right) = (2k - 1)\frac{\pi}{2}.$$

In this case formula (3.14), instead of having $2k$ constants θ_i, $i = 1, 2, \cdots, 2k$, has now only $(k - 1)$ independent constants $\theta_2, \theta_3, \cdots, \theta_k$. For $k = 1$ no such parameter appears; witness formula (5.1). We shall derive here the simplified form which formula (3.14) takes when $f_k(z)$ is typically-real.

We suppose now that $f_k(z)$ is typically-real of class k and is given by the power series

$$f_k(z) = z^k + \sum_{k+1}^{\infty} a_n^{(k)} z^n, \tag{5.3}$$

where $a_n^{(k)}$ are real numbers. If we substitute the values (5.2) in the second form of (3.14) the formula simplifies to

---

[4] See W. Rogosinski, "Über positive harmonische Entwicklungen und typisch-reelle Potenzreihen," *Mathematische Zeitschrift*, Bd. 35, (1932), pp. 93–121.

[5] See W. Rogosinski, loc. cit.

$$f_k(z) = \frac{z^k F(z)}{(1 - z^2) \cdot \prod_{j=2}^{k} (1 - 2z \cos \theta_j + z^2)} = f_1(z) \cdot \prod_{j=2}^{k} h_{\theta_j}(z). \tag{5.4}$$

We have thus shown that $f_k(z)$ can be written in the form (5.4). Conversely, suppose $F(z)$ is regular for $|z| \leqq 1$, $F(0) = 1$, $F(z)$ is real on the real axis, and $\Re F(z) > 0$ for $|z| < 1$. Then $f_k(z)$ formed as in (5.4) is typically-real of class k with respect to the unit circle.

From (5.4) it follows that

$$|f_k(re^{i\theta})| \leqq \frac{r^k}{(1 - r)^{2k}}, \tag{5.5}$$

For functions of class k which are not necessarily real on the real axis we have seen, as in (4.2), that the coefficients a_n satisfy the inequalities:

$$|a_n| \leqq \frac{n}{k} \frac{(n + k - 1)!}{(2k - 1)!\,(n - k)!} = O(n^{2k}), \qquad n > k. \tag{5.6}$$

However, for functions typically-real of class k we might naturally expect a better estimate resulting from formula (5.4). If $f_1(z)$ is denoted by

$$f_1(z) = z + \sum_{2}^{\infty} a_n^{(1)} z^n \tag{5.7}$$

then it is well-known[6] that $|a_n^{(1)}| \leqq n$. Since also

$$h_\alpha(z) \equiv z(1 - 2z \cos \alpha + z^2)^{-1} = \sum_{1}^{\infty} \frac{\sin m\alpha}{\sin \alpha} z^m$$

it follows that each $h_{\theta_j}(z)$ and $f_1(z)$ in (5.4) has the function $z(1 - z)^{-2}$ as a majorant. Hence $z^k(1 - z)^{-2k}$ is a majorant for $f_k(z)$ whence

$$|a_n^{(k)}| \leqq \frac{(n + k - 1)!}{(2k - 1)!\,(n - k)!} = O(n^{2k-1}) \tag{5.8}$$

$$\overline{\lim_{n \to \infty}} \left| \frac{a_n^{(k)}}{n^{2k-1}} \right| \leqq \frac{1}{(2k - 1)!} \tag{5.9}$$

which is clearly a better estimate than the one in (5.6), holding for the larger class of functions not necessarily real on the real axis. The equality sign is attained in (5.8) by the function $z^k(1 - z)^{-2k}$.

6. **The sub-class of odd functions of** (S_k). Let k be an odd integer and $\phi_k(z)$ an odd function of class (S_k) having the power series

$$\phi_k(z) = z^k + \sum_{n=1}^{\infty} a_{k+2n} z^{k+2n}. \tag{6.1}$$

[6] See W. Rogosinski, loc. cit.

Since $\phi_k(z)$ is an odd function it follows that

$$\theta_{k+s} = \theta_s + \pi, \qquad s = 1, 2, \cdots, k$$

(6.2)
$$\tfrac{1}{2}\sum_1^{2k} \theta_s = \tfrac{1}{2}\sum_1^k \theta_s + \frac{k\pi}{2}, \qquad k \text{ odd}.$$

Consequently, formula (3.14) now assumes the form

(6.3)
$$\phi_k(z) = \frac{z^k}{\prod_{s=1}^{k} (1 - z^2 e^{-2i\theta_s})} \cdot \left[1 + e^{-\frac{1}{2}i\sum_1^k \theta_s} \cdot \cos\left(\tfrac{1}{2}\sum_1^k \theta_s\right)\{F(z) - 1\}\right].$$

Since $\phi_k(z)$ is an odd function, $F(z)$ is an even function of z. From (6.3) we immediately obtain for the rate of growth of $\phi_k(z)$

(6.4)
$$|\phi_k(re^{i\theta})| \leqq \frac{r^k}{(1 - r^2)^k}\left(1 + \frac{2r^2}{1 - r^2}\right) = \frac{r^k(1 + r^2)}{(1 - r^2)^{k+1}}.$$

From (6.3) it also follows that $z^k(1 + z^2)/(1 - z^2)^{k+1}$ is a majorant for $\phi_k(z)$ whence we obtain

(6.5)
$$|a_{k+2n}| \leqq \frac{k + 2n}{k + n} \cdot \frac{(k + n)!}{k!\, n!} = O(n^k)$$

(6.6)
$$\overline{\lim_{n \to \infty}} \left|\frac{a_n}{n^k}\right| \leqq \frac{1}{k!\, 2^{k-1}}.$$

If, as well as being an odd function, $\phi_k(z)$ is also real on the real axis, we have symmetry about both the real and imaginary axes. In this case we have

(6.7)
$$\theta_{k-s+2} = \pi - \theta_s, \qquad s = 1, 2, \cdots, \frac{k - 1}{2}, \qquad k \text{ odd}.$$

Hence formula (5.4) reduces to

(6.8)
$$\phi_k(z) = \frac{z^k F(z)}{1 - z^2} \cdot \prod_{s=2}^{\frac{1}{2}(k+1)} (1 - 2z^2 \cos 2\theta_s + z^4)^{-1}.$$

Here again, as in (6.3), the function $z^k(1 + z^2)/(1 - z^2)^{k+1}$ is a majorant for $\phi_k(z)$. Consequently (6.5) is not improved upon through (6.8) when the coefficients a_{k+2n} are all real numbers.

7. **Analytic functions star-like in $2k$ directions.** We shall suppose that for all values of $r < 1$ and near to one,

(7.1)
$$w = g_k(z) = z + \sum_2^\infty C_n^{(k)} z^n,$$

regular for $|z| < 1$, maps $|z| = r$ into a contour C_r such that each ray arg $w = (p/k)\pi$, $p = 0, 1, 2, \cdots 2k - 1$, cuts it in eactly one point. We shall say that $g_k(z)$ is then star-like in the directions of these rays. $g_k(z)$ may or may not

be univalent for $|z| < 1$. The case $k = 1$ has already been discussed in an earlier paper by the author.[7]

Let $g_k(z) = \rho e^{i\phi}$. Then since $g_k(z)$ is star-like in the direction of the real axis we may assume that $0 \leqq \phi < 2\pi$. Suppose that $g_k(z) = u + iv$. Then

$$\mathfrak{I}[g_k(z)]^k = \rho^k \sin k\phi = \rho^k \sin\left(k \tan^{-1} \frac{v}{u}\right)$$

$$\mathfrak{I}[g_k(z)]^k = 0 \quad \text{when, and only when} \tag{7.2}$$

$$\tan^{-1}\frac{v}{u} = \frac{p\pi}{k}, \qquad p = 0, 1, 2, \cdots, 2k - 1.$$

Moreover, $\mathfrak{I}[g_k(z)]^k$ is alternatively positive and negative in the successive intervals between these points. Since there is only one value of z on $|z| = r < 1$ for each p for which (7.2) is true (on account of our hypothesis of $g_k(z)$ being star-like on the $2k$ rays) we have the following theorem:

THEOREM. *A necessary condition that a function $g_k(z)$, regular for $|z| < 1$, $g_k(0) = 0$, $g_k'(0) = 1$, be star-like in the $2k$ directions arg $w = (p/k)\pi$, $p = 0, 1, 2, \cdots, 2k - 1$, is that $[g_k(z)]^k$ belong to (S_k).*

We can therefore write

$$g_k(z) = [f_k(z)]^{1/k} \tag{7.3}$$

where $f_k(z)$ belongs to (S_k). We see from (3.14) and (7.3) that we may take as a majorant for $g_k(z)$ the function

$$\frac{z(1+z)^{1/k}}{(1-z)^{2+(1/k)}} = \sum_1^\infty d_n^{(k)} z^n. \tag{7.4}$$

Consequently we obtain the following inequalities for the coefficients of the power series for $g_k(z)$:

$$\begin{aligned} |C_n^{(k)}| \leqq d_n^{(k)} = {} & \frac{\frac{(-1)^n}{k}\left(1-\frac{1}{k}\right)\left(2-\frac{1}{k}\right)\cdots\left(n-2-\frac{1}{k}\right)}{(n-1)!} \\ & + \frac{\left(2+\frac{1}{k}\right)\left(3+\frac{1}{k}\right)\cdots\left(n+\frac{1}{k}\right)}{(n-1)!} \\ & + \sum_{p=1}^{n-2} \frac{(-1)^p\left(\frac{-1}{k}\right)\left(1-\frac{1}{k}\right)\left(2-\frac{1}{k}\right)\cdots\left(p-1-\frac{1}{k}\right)}{p!} \\ & \qquad \cdot \frac{\left(2+\frac{1}{k}\right)\left(3+\frac{1}{k}\right)\cdots\left(n-p+\frac{1}{k}\right)}{(n-p-1)!} \end{aligned} \tag{7.5}$$

In particular, $|C_n^{(1)}| \leqq n^2$.

[7] See M. S. Robertson, loc. cit.

Of special interest is the case when $g_k(z)$ is star-like along a denumerable infinity of rays. If we denote the coefficients of the power series for $g_k(z)$ in this case by $C_n^{(\infty)}$ we then have from (7.5)

$$| C_n^{(\infty)} | \leqq \lim_{k\to\infty} d_n^{(k)} = n. \tag{7.6}$$

The inequality (7.6) is well-known for functions star-like in every direction with respect to the unit circle.

From (3.14) and (7.3) we also have for the rate of growth of $| g_k(z) |$

$$| g_k(re^{i\theta}) | \leqq \frac{r(1+r)^{1/k}}{(1-r)^{2+(1/k)}}. \tag{7.7}$$

In order to show that the order of $| C_n^{(k)} |$ in (7.5) is $O(n^{1+k^{-1}})$ we may argue as follows. From the second form of the formula (3.14) together with (7.3) we may write

$$g_k(z) = \Psi_k(z)\cdot\Phi_k(z) \tag{7.8}$$

where

$$\Psi_k(z) \equiv z\cdot\prod_{s=1}^{2k}(1 - ze^{-i\theta_s})^{-1/k}$$

$$\Phi_k(z) \equiv [1 + ie^{-i\sigma_k}\cdot\sin\sigma_k\,\{F(z) - 1\}]^{1/k}$$

$$\sigma_k = \tfrac{1}{2}\sum_{1}^{2k}\theta_s$$

The function $\Psi_k(z)$ is univalent and star-like with respect to the unit circle since $\Re\,\frac{z\psi_k'(z)}{\Psi_k(z)} > 0$ for $| z | < 1$, $\Psi_k'(0) \neq 0$. If the numbers θ_s are all distinct $w = \Psi_k(z)$ maps the unit circle on the star domain whose boundary consists of $2k$ rays, *arg* w = constant, broken off at various distances from the origin. The angle between any two successive rays is π/k radians.

Since $\Psi_k(z)$ is star-like it follows[8] that

$$\frac{1}{2\pi}\int_0^{2\pi} | \Psi_k(re^{i\theta}) |\, d\theta \leqq \frac{r}{1-r^2} \tag{7.9}$$

and from (3.18) we have

$$| \Phi(re^{i\theta}) | \leqq \left(\frac{1+r}{1-r}\right)^{1/k}. \tag{7.10}$$

Since we also have

$$C_n^{(k)} = \frac{1}{2\pi r^n}\int_0^{2\pi} g_k(re^{i\theta})e^{-ni\theta}\,d\theta \tag{7.11}$$

---

[8] See M. S. Robertson, "On the theory of univalent functions," *Annals of Mathematics*, Vol. 37, (1936), pp. 374–408, esp. p. 389.

we obtain, using (7.8), (7.9), and (7.10)

$$
(7.12)\qquad
\begin{aligned}
|C_n^{(k)}| &\leqq \frac{1}{2\pi r^n}\cdot\int_0^{2\pi} |\Psi_k(re^{i\theta})|\cdot|\Phi_k(re^{i\theta})|\,d\theta \\
&\leqq \frac{(1+r)^{1/k}}{r^n(1-r)^{1/k}}\cdot\frac{1}{2\pi}\int_0^{2\pi}|\Psi_k(re^{i\theta})|\,d\theta \\
&\leqq r^{-n}\left(\frac{1+r}{1-r}\right)^{1/k}\cdot\frac{r}{1-r^2} \\
&\leqq \frac{1}{r^{n-1}(1+r)^{1-(1/k)}(1-r)^{1+(1/k)}}.
\end{aligned}
$$

Taking $r = 1 - 1/n$ we have

$$
(7.13)\qquad |C_n^{(k)}| \leqq \frac{n^{1+(1/k)}}{\left(1-\frac{1}{n}\right)^{n-1}\left(2-\frac{1}{n}\right)^{1-(1/k)}} \leqq \frac{\left(1+\frac{1}{n-1}\right)^{n-1}\cdot n^{1+(1/k)}}{\left(2-\frac{1}{n}\right)^{1-(1/k)}} < en^{1+(1/k)}.
$$

8. **The multivalency of functions of** (S_k). An analytic function $f(z)$ which is holomorphic for $|z| < R$ and which takes on no value more than p times in this circle and which does take on at least one value p times in this circle is said to be multivalent of order p with respect to the circle $|z| = R$. With regard to the functions $f_k(z)$ of class (S_k) considered above one cannot say in general that $f_k(z)$ is multivalent of order k, but merely that if $f_k(z)$ is multivalent of order p then $p \geqq k$. However, one can say more about the class of functions $F_k(z)$ defined as

$$
(8.1)\qquad F_k(z) \equiv k\int_0^z \frac{f_k(z)}{z}\,dz = z^k + k\sum_{k+1}^{\infty}\frac{a_n}{n}\cdot z^n.
$$

On account of the identity

$$
(8.2)\qquad \Im\{kf_k(z)\} = \Im\{zF_k'(z)\} = -\frac{\partial u(r,\theta)}{\partial\theta}
$$

where $\Im$ denotes "imaginary part of" and where $\Re\, F_k(re^{i\theta}) = u(r,\theta)$ it follows that each circle $|z| = r < 1$ is mapped into a contour consisting of k loops about the origin, each separate loop having the property that no straight line parallel to the imaginary axis cuts it in more than two points. Hence for any z_0 in the unit circle the point $w_0 = F_k(z_0)$ in the w-plane can lie within at most k loops about it. Hence no value is taken on by $F_k(z)$ more than k times. It follows then that the functions $F_k(z)$ are multivalent of order k. In particular if $k = 1$, $F_1(z)$ is a univalent function, even though $f_1(z)$ is not.

For the functions $g_k(z)$ considered in section seven, since $g_k(z)$ is star-like in

the directions of $2k$ rays from the origin, $2k$-wise symmetric, the functions $\mu_k(z)$ defined as

$$\mu_k(z) \equiv \int_0^z \frac{g_k(z)}{z}\,dz = z + \sum_2^\infty \frac{C_n^{(k)}}{n} z^n \tag{8.3}$$

are, as we have shown elsewhere,[9] not only univalent, but are "convex in k directions." Thus $w = \mu_k(z)$ maps each circle $|z| = r < 1$ into a simple closed contour having the property that no straight line parallel to any one of the k directions $\arg w = (p/k)\pi + \frac{1}{2}\pi$, $p = 0, 1, 2, \cdots, 2k - 1$, cuts the contour in more than two points. If one denotes the coefficients of the power series for $\mu_k(z)$ by $\beta_n^{(k)}$ one has from (7.5) and (7.13) and from the fact that $n\beta_n^{(k)} = C_n^{(k)}$ the inequalities

$$|\beta_n^{(k)}| \leqq A(k)n^{1/k} < en^{1/k} \qquad \text{for all } n, \tag{8.4}$$

where $A(k)$ is a constant depending upon k and not upon n. In particular, if $\mu_k(z)$ is convex in every direction one obtains the well-known inequality $|\beta_n^{(\infty)}| \leqq 1$ as is seen from (7.5).

YALE UNIVERSITY.

[9] See M. S. Robertson, loc. cit.

ANNALS OF MATHEMATICS
Vol. 38, No. 4, October, 1937

ÜBER DIE PARSEVALSCHE GLEICHUNG FÜR VERALLGEMEINERTE FOURIERSCHE INTEGRALE

BY B. LEWITAN

(Received April 26, 1937)

1. Es sei $f(x)$ eine für $-\infty < x < \infty$ erklärte komplexwertige Funktion für welche die Funktion $[f(x)]^2/(1 + x^2)$ in $[-\infty, \infty]$ absolut integrierbar ist, also eine Funktion der Klasse F_1^2 nach S. Bochner.

Bekanntlich kann man der Funktion $f(x)$ ihre sogenannte verallgemeinerte Fouriersche Transformierte

$$U(\alpha) = \underset{A\to\infty}{\text{l.i.m.}}\ \frac{1}{2\pi}\int_{-A}^{A} f(x)\,\frac{e^{-i\alpha x} - L_1(\alpha x)}{-ix}\,dx$$

wo

$$L_1(\alpha x) = \begin{cases} 1 & |x| \leqq 1 \\ 0 & |x| > 1 \end{cases}$$

zuordnen.

SATZ 1. *Ist*

$$U(\alpha) = \underset{A\to\infty}{\text{l.i.m.}}\ \frac{1}{2\pi}\int_{-A}^{A} f(x)\,\frac{e^{-i\alpha x} - L_1(\alpha x)}{-ix}\,dx$$

$$V(\alpha) = \underset{A\to\infty}{\text{l.i.m.}}\ \frac{1}{2\pi}\int_{-A}^{A} g(x)\,\frac{e^{-i\alpha x} - L_1(\alpha x)}{-ix}\,dx$$

$$E(\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\,g(x)\,\frac{e^{-i\alpha x} - L_2(\alpha x)}{-x^2}\,dx$$

$$L_2(\alpha x) = \begin{cases} 1 - i\alpha x & |x| \leqq 1 \\ 0 & |x| > 1 \end{cases}$$

wo $f(x), g(x) \in F_1^2$, so besteht für jedes $\epsilon > 0$ die Beziehung

$$\frac{\Delta_{2\epsilon}^2 E(\alpha)}{2\epsilon} = \frac{E(\alpha + 2\epsilon) - 2E(\alpha) + E(\alpha - 2\epsilon)}{2\epsilon}$$

$$= \frac{1}{2\epsilon}\int_{-\infty}^{\infty} \{U(\alpha - \beta + \epsilon) - U(\alpha - \beta - \epsilon)\}\cdot\{V(\beta + \epsilon) - V(\beta - \epsilon)\}\,d\beta.$$

BEWEIS. Aus

$$\frac{U(\alpha + \epsilon) - U(\alpha - \epsilon)}{2} = \underset{A\to\infty}{\text{l.i.m.}}\ \frac{1}{2\pi}\int_{-A}^{A} f(x)\,\frac{\sin \epsilon x}{x}\,e^{-i\alpha x}\,dx$$

$$\frac{V(\alpha + \epsilon) - V(\alpha - \epsilon)}{2} = \underset{A\to\infty}{\text{l.i.m.}}\ \frac{1}{2\pi}\int_{-A}^{A} g(x)\,\frac{\sin \epsilon x}{x}\,e^{-i\alpha x}\,dx$$

folgt nach der gewöhnlichen Parsevalschen Gleichung

$$\frac{1}{4\epsilon}\int_{-\infty}^{\infty}\{U(\alpha-\beta+\epsilon)-U(\alpha-\beta-\epsilon)\}\cdot\{V(\beta+\epsilon)-V(\beta-\epsilon)\}d\beta$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}f(x)g(x)\,\frac{\sin^2\epsilon x}{\epsilon x^2}\,e^{-i\alpha x}dx=\frac{\Delta_{2\epsilon}^2E(\alpha)}{4\epsilon},$$

was zu beweisen war.

2. SATZ 2. *Damit eine Funktion $F(x)$ gleichmässig in $[-\infty, \infty]$ und beliebig genau durch Integrale $\int_{-\infty}^{\infty}\varphi(\alpha)e^{i\alpha x}d\alpha$ mit $\int_{-\infty}^{\infty}|\varphi(\alpha)|\,d\alpha<\infty$ approximierbar sei ist notwendig und hinreichend, dass $F(x)$ stetig ist und für $|x|\to\infty$ nach Null strebt.*

BEWEIS. Es sei gegeben, dass für jedes $\epsilon>0$ eine Funktion $\varphi_\epsilon(\alpha)$ existiert, für welche

$$\sup_{-\infty<x<\infty}|F(x)-\int_{-\infty}^{\infty}e^{i\alpha x}\varphi_\epsilon(\alpha)d\alpha|<\epsilon,\qquad\int_{-\infty}^{\infty}|\varphi_\epsilon(\alpha)|\,d\alpha<\infty.$$

Bekanntlich ist $\int_{-\infty}^{\infty}e^{i\alpha x}\varphi_\epsilon(\alpha)d\alpha$ stetig und strebt nach Null für $|x|\to\infty$ woraus die Notwendigkeit unmittelbar folgt.

Es sei jetzt $F(x)$ als stetig vorausgesetzt und es möge die Beziehung $|F(x)|<\epsilon$ für $|x|>|x_\epsilon|$ statt finden, wo $\epsilon>0$ beliebig gewählt werden kann. Wir nehmen die Funktion

$$\lambda_\epsilon(x)=\begin{cases}1 & |x|\leqq x_\epsilon\\ -|x|+x_\epsilon+1 & x_\epsilon\leqq|x|\leqq x_\epsilon+1\\ 0 & |x|\geqq x_\epsilon+1\end{cases}$$

und setzen $F_\epsilon(x)=\lambda_\epsilon(x).\,F(x)$. Offenbar ist $|F(x)-F_\epsilon(x)|<\epsilon$ für $-\infty<x<\infty$. Die Funktion $F_\epsilon(x)$ ist stetig und gleich Null ausserhalb des Intervalls $(-x_\epsilon-1, x_\epsilon+1)$. Daher kann man $F_\epsilon(x)$ gleichmässig durch trigonometrische Integrale approximieren, woraus die Hinlänglichkeit folgt.

3. Die den Bedingungen des Satzes 2 genügende Funktion $F(x)$ gehört der Klasse F_1^2. Wir wollen zeigen, wie man vermöge der verallgemeinerten Fourierschen Transformierten $U(\alpha)$ für $F(x)$ diese Funktion konstruieren kann. Es genügt dazu zu zeigen, wie vermöge der Funktion $U(\alpha)$ die gewöhnliche Fouriersche Transformierte für $F_\epsilon(x)$ (und somit auch $F_\epsilon(x)$ selbst) gefunden werden kann. Ist

$$U(\alpha)=\underset{A\to\infty}{\mathrm{l.i.m.}}\,\frac{1}{2\pi}\int_{-A}^{A}f(x)\,\frac{e^{-i\alpha x}-L_1(\alpha x)}{-ix}\,dx$$

$$V(\alpha)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\lambda_\epsilon(x)\,\frac{e^{-i\alpha x}-L_1(\alpha x)}{-ix}\,dx=\frac{2}{\pi}\int_0^{\alpha}\frac{\sin(x_\epsilon+\frac{1}{2})\,\alpha\sin\frac{1}{2}\alpha}{\alpha^2}\,dx+C$$

so folgt nach dem Satze 1

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} F_\epsilon(x)\, e^{-i\alpha x}\, dx = \lim_{\eta\to 0} \frac{1}{2\pi}\int_{-\infty}^{\infty} F_\epsilon(x)\, \frac{\sin^2 \eta x}{(\eta x)^2}\, e^{-i\alpha x}\, dx$$

$$= \lim_{\eta\to 0} \int_{-\infty}^{\infty} U(\alpha-\beta)\, \frac{\Delta_{2\eta}^2 V(\beta)}{4\eta^2}\, d\beta = \int_{-\infty}^{\infty} U(\alpha-\beta)\, V''(\beta)\, d\beta = \int_{-\infty}^{\infty} U(\alpha-\beta)$$

$$\frac{\{x_\epsilon \cos (x_\epsilon + \tfrac{1}{2})\, \beta \sin \tfrac{1}{2}\beta + \tfrac{1}{2}\sin (x_\epsilon + 1)\beta\}\, \beta - 2 \sin (x_\epsilon + \tfrac{1}{2})\, \beta \sin \tfrac{1}{2}\beta}{\beta^3}\, d\beta$$

wobei die Möglichkeit des letzten Grenzüberganges leicht bestetigt werden kann.

CHARKOW, MATHEMATISCHES INSTITUT AN DER UNIVERSITÄT A. M. GORKI.

ANNALS OF MATHEMATICS
Vol. 38, No. 4, October, 1937

ON CERTAIN METRIC SPACES ARISING FROM EUCLIDEAN SPACES BY A CHANGE OF METRIC AND THEIR IMBEDDING IN HILBERT SPACE[1]

BY I. J. SCHOENBERG

(Received January 7, 1937)

1. W. A. Wilson ([9])[2] has recently investigated those metric spaces which arise from a metric space by taking as its new metric a suitable (one variable) function of the old one. He considered in particular the euclidean straight line R_1 whose metric $\delta = \overline{PP'}$ is changed to $\Delta = d(P, P') = \overline{PP'}^{\frac{1}{2}}$ and showed that this new metric space can be imbedded[3] in Hilbert space $\mathfrak{H}$. Here the old metric δ and the new metric Δ are connected by the relation $\Delta^2 = \delta$.

In an article soon to appear ([5]), John von Neumann and the author have determined all the functions $f(\delta)$ such that if R_1 is provided with the new metric Δ, defined by $\Delta^2 = f(\delta)$, $\delta = \overline{PP'}$, the new metric space thus arising shall be imbeddable in $\mathfrak{H}$. They are of the form

$$f(\delta) = \int_0^\infty \frac{\sin^2 (s\delta)}{s^2} \, d\alpha(s), \tag{1}$$

where $\alpha(s)$ is non-decreasing for $0 \leqq s < \infty$ and such that $\int_1^\infty s^{-2}\, d\alpha(s)$ exists.

Wilson's case $f(\delta) = \delta$ is included in the general formula on account of

$$\delta = \frac{2}{\pi} \int_0^\infty \frac{\sin^2 (s\delta)}{s^2} \, ds, \qquad (\delta \geqq 0). \tag{2}$$

In the present note Wilson's example is extended to higher dimensional euclidean spaces, its chief result being the following theorem.

THEOREM 1. *If we change the metric of the euclidean space R_m from the euclidean distance $\overline{PP'}$ to the new distance*

$$d(P, P') = \overline{PP'}^{\gamma}, \qquad (0 < \gamma < 1),^{4} \tag{3}$$

the new space $\mathfrak{R}_m^{(\gamma)}$ thus arising may be imbedded isometrically in the Hilbert space $\mathfrak{H}$.

---

[1] Presented to the American Mathematical Society, February 20, 1937.

[2] The numbers in square brackets refer to the list of references at the end of this note.

[3] Here and throughout this note the word *imbedding* is meant in the sense of *isometrical imbedding*.

[4] The case $\gamma = 1$ is trivial. The theorem does not hold for $\gamma = 0$, for the space $\mathfrak{R}_m^{(0)}$, with $d(P, P') = 1$ if $P \neq P'$ and $d(P, P) = 0$, is obviously not separable. The constant 1 is the best constant, for $\mathfrak{R}_m^{(\gamma)}$ is not a metric space if $\gamma > 1$.

2. As for all such imbedding problems into $\mathfrak{H}$, the proof of Theorem 1 is based on the following theorem of Menger ([4]).

A metric space $\mathfrak{R}$ can be imbedded in $\mathfrak{H}$ if and only if $\mathfrak{R}$ is separable and every set of $n + 1$ ($n = 2, 3, 4, \cdots$) distinct points of $\mathfrak{R}$ can be imbedded in R_n.

Therefore, as our $\mathfrak{R}_m^{(\gamma)}$ is obviously separable, it suffices to show that any $n + 1$ distinct points $P_0, P_1, \cdots, P_n$, of $\mathfrak{R}_m^{(\gamma)}$ can be imbedded in R_n, i.e., there exist $n + 1$ points $Q_0, Q_1, \cdots, Q_n$, of R_n such that $\overline{Q_\mu Q_\nu} = \overline{P_\mu P_\nu}^\gamma$ ($\mu, \nu = 0, 1, \cdots, n$).[5] This "finite" imbedding problem is readily solved by means of the following theorem ([6], Theorem 1, p. 724).

The quantities $a_{\mu\nu}$ ($\mu, \nu = 0, 1, \cdots, n$; $a_{\mu\nu} = a_{\nu\mu} > 0$ if $\mu \neq \nu$, $a_{\mu\mu} = 0$) are the distances of $n + 1$ points $Q_0, Q_1, \cdots, Q_n$, of R_n, i.e. $a_{\mu\nu} = \overline{Q_\mu Q_\nu}$, if and only if the quadratic form

$$F(x, x) = \tfrac{1}{2} \sum_{j,k=1}^{n} (a_{0j}^2 + a_{0k}^2 - a_{jk}^2) x_j x_k$$

is positive, i.e. always $\geqq 0$. If this form is positive definite, the points Q_μ are the vertices of a n-simplex.[6]

Our finite imbedding problem $a_{\mu\nu} = \overline{P_\mu P_\nu}^\gamma = Q_\mu Q_\nu$ is therefore contained (for $\alpha = 2\gamma$) in the following theorem.

THEOREM 2. *If $P_0, P_1, \cdots, P_n$, are distinct points of a euclidean space R_m ($m \geqq 1$), the quadratic form*

$$F^{(\alpha)}(x, x) = \tfrac{1}{2} \sum_{j,k=1}^{n} (\overline{P_0 P_j}^\alpha + \overline{P_0 P_k}^\alpha - \overline{P_j P_k}^\alpha) x_j x_k \quad (0 < \alpha < 2) \tag{4}$$

is positive definite.[7]

Note that in order to prove Theorem 1 we need only to know that $F^{(\alpha)}(x, x) \geqq 0$. Its positive definiteness means that in order to imbed into $\mathfrak{H}$ any $n + 1$ distinct points of $\mathfrak{R}_m^{(\gamma)}$ we need fully all dimensions of a n-dim. subspace of $\mathfrak{H}$, i.e. a R_n.

[5] L. M. Blumenthal ([2], Corollary, p. 402) proved the following result. *If P_i($i = 0, 1, 2, 3$) are four points of a metric space $\mathfrak{R}$, for any nonnegative number γ, not exceeding $\frac{1}{2}$, there exist four points Q_i($i = 0, 1, 2, 3$) of R_3 such that $\overline{Q_i Q_j} = \{d(P_i, P_j)\}^\gamma$($i, j = 0, 1, 2, 3$).*

This result is *not* contained in our present problem, for the distances $d(P_i, P_j)$ are not assumed in Blumenthal's theorem to be the edges of a euclidean tetrahedron. If this assumption is added, as for instance by assuming $\mathfrak{R}$ to be a euclidean space, Blumenthal conjectures that the inequality $0 \leqq \gamma \leqq \frac{1}{2}$ of his theorem may be replaced by $0 \leqq \gamma \leqq 1$ (loc. cit., concluding remark of section 4, p. 403). Theorem 2 below proves this conjecture and extends it from four points to $n + 1$ points.

[6] This elementary theorem is in substance identical with the well known correspondence between lattices of points and positive definite quadratic forms. See H. Minkowski, *Gesammelte Abhandlungen*, vol. 1, pp. 243–254, where also references to Gauss and Dirichlet are found. For an imbedding problem of arithmetical nature see I. J. Schoenberg, [7].

[7] Communicating the proof of Theorem 2 to Prof. G. Szegö, my letter and one of his crossed each other; in his letter Prof. Szegö proves independently and in a different way Theorem 2 for $\alpha = 1$ and $m = 1, 2$ and 3. An extension of his proof to arbitrary α ($0 < \alpha < 2$) is obvious, but not an extension to all dimensions m.

3. Let us pass now to the proof of Theorem 2. Although this theorem is algebraic in nature, at least for rational values of α, an algebraic proof would probably be difficult and complicated. The following proof is elementary but uses transcendental means. To simplify notations we prove it first for $m = 3$, i.e., $P_0, P_1, \cdots, P_n$, are points in ordinary 3-space.

Consider the following function of three real variables

$$\Omega(u, v, w) = \frac{1}{4\pi} \iint\limits_{\xi^2+\eta^2+\zeta^2=1} e^{i(u\xi+v\eta+w\zeta)}\, d\sigma = \mathfrak{M}\{e^{i(u\xi+v\eta+w\zeta)}\},$$

which is the mean value of the function $e^{i(u\xi+v\eta+w\zeta)}$ over the spherical shell $\xi^2 + \eta^2 + \zeta^2 = 1$. $\Omega(u, v, w)$ is obviously invariant with respect to rigid rotations around the origin and is therefore a function of $r = (u^2 + v^2 + w^2)^{\frac{1}{2}}$ only, which we denote by $\Omega(r)$. Now

$$\Omega(r) = \Omega(0, 0, r) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} e^{ir\cos\theta} \sin\theta\, d\varphi\, d\theta = \tfrac{1}{2} \int_0^{\pi} e^{ir\cos\theta} \sin\theta\, d\theta = \frac{\sin r}{r},$$

hence

$$\Omega(r) = \frac{\sin r}{r} = \mathfrak{M}\{e^{i(u\xi+v\eta+w\zeta)}\} \quad (r = (u^2 + v^2 + w^2)^{\frac{1}{2}}). \tag{5}$$

Let $P_0 = (0, 0, 0)$, $P_i = (u_i, v_i, w_i)$, indicate the coordinates of our points in R_3. For $s \geqq 0$ we have

$$\begin{aligned}\Omega(s \cdot \overline{P_\mu P_\nu}) &= \Omega(s \cdot \sqrt{(u_\mu - u_\nu)^2 + (v_\mu - v_\nu)^2 + (w_\mu - w_\nu)^2}) \\ &= \mathfrak{M}\{e^{is[(u_\mu-u_\nu)\xi+(v_\mu-v_\nu)\eta+(w_\mu-w_\nu)\zeta]}\},\end{aligned}$$

whence

$$\Omega(s \cdot \overline{P_\mu P_\nu}) = \mathfrak{M}\{e^{is(u_\mu\xi+v_\mu\eta+w_\mu\zeta)} \cdot e^{-is(u_\nu\xi+v_\nu\eta+w_\nu\zeta)}\} \qquad (s \geqq 0). \tag{6}$$

On the other hand we have (as is readily seen by substituting st^{-1} for s in the integral) for $0 < \alpha < 2$

$$t^\alpha = c(\alpha) \cdot \int_0^\infty \frac{1 - \Omega(ts)}{s^2} s^{1-\alpha} ds = c(\alpha) \cdot \int_0^\infty \{1 - \Omega(ts)\} s^{-1-\alpha} ds \qquad (0 < \alpha < 2; t > 0), \tag{7}$$

where

$$c(\alpha) = 1 \Big/ \int_0^\infty \{1 - \Omega(s)\} s^{-1-\alpha} ds \qquad (0 < \alpha < 2).$$

We may now express our hermitian form $F^{(\alpha)}(x, \bar{x})$ as follows

$$\begin{aligned}&\tfrac{1}{2} \sum_{j,k=1}^{n} \{\overline{P_0 P_j}^\alpha + \overline{P_0 P_k}^\alpha - \overline{P_j P_k}^\alpha\} x_j \bar{x}_k \\ &= \frac{c(\alpha)}{2} \int_0^\infty s^{-1-\alpha} \sum_1^n \{1 - \Omega(s \cdot \overline{P_0 P_j}) - \Omega(s \cdot \overline{P_0 P_k}) + \Omega(s \cdot \overline{P_j P_k})\} x_j \bar{x}_k \cdot ds,\end{aligned} \tag{8}$$

where $x_1, \cdots, x_n$, are arbitrary complex numbers. Writing

$$x_0 = -\sum_1^n x_j$$

we have

$$\sum_1^n \{1 - \Omega(s\cdot\overline{P_0P_j}) - \Omega(s\cdot\overline{P_0P_k}) + \Omega(s\cdot\overline{P_jP_k})\}x_j\bar{x}_k$$

$$= \left|\sum_1^n x_j\right|^2 - \sum_1^n \Omega(s\cdot\overline{P_0P_j})x_j\cdot\sum_1^n \bar{x}_k - \sum_1^n \Omega(s\cdot\overline{P_0P_k})\bar{x}_k\cdot\sum_1^n x_j + \sum_1^n \Omega(s\cdot\overline{P_jP_k})x_j\bar{x}_k$$

$$= |x_0|^2 + \bar{x}_0\cdot\sum_1^n \Omega(s\cdot\overline{P_0P_j})x_j + x_0\cdot\sum_1^n \Omega(s\cdot\overline{P_0P_k})\bar{x}_k + \sum_1^n \Omega(s\cdot\overline{P_jP_k})x_j\bar{x}_k$$

$$= \sum_{\mu,\nu=0}^n \Omega(s\cdot\overline{P_\mu P_\nu})x_\mu\bar{x}_\nu,$$

which, in view of (6), is equal to

$$\mathfrak{M}\left\{\left| x_0 + \sum_{j=1}^n x_j e^{is(u_j\xi+v_j\eta+w_j\zeta)}\right|^2\right\}.$$

Now (8) becomes

$$(9)\quad \begin{aligned}\tfrac{1}{2}\sum_{j,k=1}^n (\overline{P_0P_j}^\alpha + \overline{P_0P_k}^\alpha - \overline{P_jP_k}^\alpha)x_j\bar{x}_k \\ = \frac{c(\alpha)}{2}\cdot\int_0^\infty s^{-1-\alpha}\cdot\mathfrak{M}\left\{\left|x_0 + \sum_{j=1}^n x_j e^{is(u_j\xi+v_j\eta+w_j\zeta)}\right|^2\right\}ds \geqq 0.\end{aligned}$$

Here we have the equality sign if and only if

$$(10)\qquad x_0 + \sum_{j=1}^n x_j e^{is(u_j\xi+v_j\eta+w_j\zeta)} = 0$$

holds identically in s and the direction cosines ξ, η, ζ. As the points (u_j, v_j, w_j) are all different and none is at the origin, a direction (ξ, η, ζ) can be found for which the inner products $u_j\xi + v_j\eta + w_j\zeta$ ($j = 1, \cdots, n$) are all different and none is zero, and now (10) implies that $x_1, \cdots, x_n$ must all vanish. Theorem 2 is thus completely proved for $m = 3$, hence also for $m = 1$ and $m = 2$.

4. An extension of the proof to any value of m is now obvious. All we have to do is to repeat the above argument with the function

$$(11)\qquad \Omega_m(r) = \mathfrak{M}\{e^{i(u_1\xi_1+\cdots+u_m\xi_m)}\},\qquad r = (u_1^2 + \cdots + u_m^2)^{\frac{1}{2}},$$

that is, the mean value of $e^{i(u_1\xi_1+\cdots+u_m\xi_m)}$ over the spherical shell

$$\xi_1^2 + \cdots + \xi_m^2 = 1.$$

Thus for $m = 1$ we have

$$\Omega_1(r) = \tfrac{1}{2}(e^{ir} + e^{-ir}) = \cos r$$

and for $m = 2$

$$\Omega_2(r) = \frac{1}{2\pi}\int_0^{2\pi} e^{i(u_1\cos\theta + u_2\sin\theta)}\,d\theta = J_0(r)$$

where $J_0(r)$ is the Bessel function.

To settle the matter there remain only two essential details to be checked. First, that the improper integral in the formula

$$(12) \qquad t^\alpha = c_m(\alpha)\int_0^\infty \frac{1-\Omega_m(ts)}{s^2}\,s^{1-\alpha}\,ds \qquad (0 < \alpha < 2; t > 0),$$

which is the analogue of (7), actually converges for $0 < \alpha < 2$. Second, that the factor

$$c_m(\alpha) = 1 \Big/ \int_0^\infty \{1 - \Omega_m(s)\} s^{-1-\alpha}\,ds$$

on the right side of (12) is defined and positive for $0 < \alpha < 2$. Both facts were obvious in the case of (7), for $\Omega_3(r) = \sin r/r$ enjoys the properties

$$-1 \leqq \Omega_3(r) \leqq 1, \qquad \Omega_3(r) = 1 - \tfrac{1}{6}r^2 + \cdots.$$

In order to establish similar properties of $Q_m(r)$, we remark that by m-dimensional polar coordinates we readily find

$$(13) \qquad \Omega_m(r) = \int_0^\pi e^{ir\cos\theta}\sin^{m-2}\theta\,d\theta \Big/ \int_0^\pi \sin^{m-2}\theta\,d\theta.$$

Expansion of the exponential integrand in its power series shows that $\Omega_m(r)$ is a real and even entire function of r. The remark that $|\Omega_m(r)| \leqq \Omega_m(0) = 1$ completes the argument.

Incidentally, (13) gives the expansion

$$(14) \quad \Omega_m(r) = 1 - \frac{r^2}{2\cdot m} + \frac{r^4}{2\cdot 4\cdot m(m+2)} - \frac{r^6}{2\cdot 4\cdot 6\cdot m(m+2)(m+4)} + \cdots,$$

from which the following expression in terms of Bessel functions of the first kind becomes apparent

$$(15) \qquad \Omega_m(r) = \Gamma\left(\frac{m}{2}\right)\left(\frac{2}{r}\right)^{\frac{1}{2}(m-2)} J_{\frac{1}{2}(m-2)}(r) \qquad (m = 1, 2, 3, \cdots).$$

5. To John von Neumann are due the following consequences of our previous results.

THEOREM 3. *Let $\mathfrak{H}_\gamma$ denote the metric space obtained from the Hilbert space $\mathfrak{H}$ by replacing its metric $d(x, x') = \| x - x' \|$ by $d_\gamma(x, x') = \| x - x' \|^\gamma$ $(0 < \gamma \leqq 1)$.*

1. *$\mathfrak{H}_\gamma$ may be imbedded in $\mathfrak{H}$.*
2. *If $0 < \gamma \leqq \delta \leqq 1$, $\mathfrak{H}_\gamma$ may be imbedded in $\mathfrak{H}_\delta$.*

The first statement is essentially equivalent to Theorem 1 on the basis of

Menger's theorem. Indeed, as any $n + 1$ points of $\mathfrak{H}_\gamma$ may be imbedded in $\mathfrak{H}$, the whole of $\mathfrak{H}_\gamma$ may thus be imbedded. This result may be stated analytically as follows: There exists a function $\varphi_\gamma(x)$, defined for all $x \in \mathfrak{H}$, with $\varphi_\gamma(x) \in \mathfrak{H}$, such that

$$\| \varphi_\gamma(x) - \varphi_\gamma(x') \| = \| x - x' \|^\gamma \qquad (x \in \mathfrak{H}, x' \in \mathfrak{H}).$$

As for the second statement, which contains the first as a special case ($\delta = 1$), consider $\mathfrak{H}_{\gamma/\delta}$. A function $\varphi_{\gamma/\delta}(x)$ which performs its imbedding in $\mathfrak{H}$ satisfies, as just mentioned, the identity

$$\| \varphi_{\gamma/\delta}(x) - \varphi_{\gamma/\delta}(x') \| = \| x - x' \|^{\gamma/\delta} \qquad (x \in \mathfrak{H}, x' \in \mathfrak{H}),$$

whence

$$\| \varphi_{\gamma/\delta}(x) - \varphi_{\gamma/\delta}(x') \|^\delta = \| x - x' \|^\gamma \qquad (x \in \mathfrak{H}, x' \in \mathfrak{H}).$$

Hence $\varphi_{\gamma/\delta}(x)$ $(x \in \mathfrak{H})$, which imbeds $\mathfrak{H}_{\gamma/\delta}$ in $\mathfrak{H}$, at the same time performs the imbedding of $\mathfrak{H}_\gamma$ in $\mathfrak{H}_\delta$.

6. Let us finally state and prove the following corollary of Theorem 2.

THEOREM 4. *If $P_0, P_1, \cdots, P_n$ are distinct points of the euclidean space R_n, the quadratic form*

$$\sum_{i,k=0}^{n} \overline{P_iP_k}^\alpha x_i x_k \qquad (0 < \alpha < 2)$$

is non-singular and its canonical representation contains one positive and n negative squares.

The sole difficulty consists in proving that the determinant

$$\det \| \overline{P_iP_k}^\alpha \|_{0,n} \neq 0.$$

Let us show that

$$\operatorname{sgn} \det \| \overline{P_iP_k}^\alpha \|_{0,n} = (-1)^n \qquad (n \geqq 1). \tag{16}$$

Now perform in R_n an ordinary inversion by reciprocal radii with respect to the sphere of center P_0 and radius $r = 1$, and let $Q_1, \cdots, Q_n$ be the transforms of the points $P_1, \cdots, P_n$ by this inversion.[8] Consider the determinant of order $n + 1$

$$D = \begin{vmatrix} 0 & 1 \\ 1 & \overline{Q_iQ_k}^\alpha \end{vmatrix} \qquad (i, k = 1, 2, \cdots, n).$$

On the one hand we find (compare Blumenthal [1], p. 424) by suitable subtractions of lines and columns

$$D = (-1)^n \det \| \overline{Q_1Q_r}^\alpha + \overline{Q_1Q_s}^\alpha - \overline{Q_rQ_s}^\alpha \| \qquad (r, s = 2, 3, \cdots, n),$$

[8] This inversion was suggested by the equivalence under inversion between the triangle inequality and Ptolemy's inequality of elementary geometry. See J. Hadamard [3], pp. 228–229.

hence

$$\operatorname{sgn} D = (-1)^n \tag{17}$$

by Theorem 2. On the other hand we have (since $r = 1$) by an elementary property of inversion

$$\overline{Q_iQ_k} = \overline{P_iP_k}/(\overline{P_0P_k}\cdot\overline{P_0P_i}), \qquad (i, k = 1, \cdots, n),$$

whence

$$D = \begin{vmatrix} 0 & 1 \\ 1 & \overline{P_iP_k}^{\alpha}/\overline{P_0P_i}^{\alpha}\cdot\overline{P_0P_k}^{\alpha} \end{vmatrix} = (\overline{P_0P_1}\cdots\overline{P_0P_n})^{-2\alpha}\cdot\begin{vmatrix} 0 & \overline{P_0P_k}^{\alpha} \\ \overline{P_0P_i}^{\alpha} & \overline{P_iP_k}^{\alpha} \end{vmatrix}.$$

Hence (17) implies (16). The theorem now follows from the classical theory of the signature of quadratic forms.

A special case of Theorem 4, where $\alpha = 1$ and $P_0, \cdots, P_n$ are equidistant points on a straight line, was discussed from the point of view of Toeplitz matrices by G. Szegö [8].

REFERENCES

(1) L. M. Blumenthal, *A note on the four point property*, Bull. Amer. Math. Soc., vol. 39 (1933), pp. 423–426.

(2) —— *New theorems and methods in determinant theory*, Duke Math. J., vol. 2 (1936), pp. 396–404.

(3) J. Hadamard, *Leçons de géométrie élémentaire*, vol. 1, Paris, 1931.

(4) K. Menger, *Die Metrik des Hilbertschen Raumes*, Anzeiger de Akad. der Wissenschaften in Wien, Math. Nat. Kl., vol. 65 (1928), pp. 159–160.

(5) John von Neumann and I. J. Schoenberg, *Fourier integrals and metric geometry*, to appear in the Transactions of the Amer. Math. Soc.

(6) I. J. Schoenberg, *Remarks to Maurice Fréchet's article . . .*, these Annals, vol. 36 (1935), pp. 724–732.

(7) —— *Regular simplices and quadratic forms*, J. London Math. Soc., vol. 12 (1937), pp. 48–55.

(8) G. Szegö, *Solutions to problem 3705 (proposed by Raphael Robinson)*, Amer. Math. Monthly, vol. 43 (1936), pp. 246–259.

(9) W. A. Wilson, *On certain types of continuous transformations of metric spaces*, Amer. J. of Math., vol. 57 (1935), pp. 62–68.

Colby College, Waterville, Maine.

ANNALS OF MATHEMATICS
Vol. 38, No. 4, October, 1937

RIEMANNIAN SPACES OF CLASS GREATER THAN UNITY

BY LUTHER PFAHLER EISENHART

(Received May 20, 1937)

The geometry of a Riemannian space V_n as a subvariety of a Riemannian space V_{n+p} was studied first by Voss.[1] In presenting this theory by means of tensor calculus Ricci[2] made use of n linearly independent mutually orthogonal unit vectors of V_{n+p} normal to V_n. This method was followed in my presentation of the theory for spaces whose fundamental quadratic form is definite or indefinite as the case may be, when V_{n+p} is a general space and in particular when it is flat.[3] Burstin and Mayer, together and individually, have developed a theory for the case when the fundamental form is definite in accordance with which the n normal vectors are chosen in a particular manner and arranged in groups, each group being treated as a unit, resulting in a generalization to varieties of higher order of the Frenet formulas for a curve.[4]

In the present paper a study is made of a V_n of class $p(> 1)$, when the fundamental form is definite or indefinite, and the vectors normal to V_n are chosen in groups somewhat after the manner of Burstin and Mayer, but the groups are not treated as units.

1. Let V_n be a Riemannian space of coordinates x^i with the fundamental tensor g_{ij}, not necessarily positive definite, and such that the Riemannian curvature tensor R_{hijk} is not zero. We assume that V_n is immersed in a flat space S_{n+p} of cartesian coordinates z^α, and that this is the flat space of lowest order in which it can be immersed, that is, V_n is of class p. If the fundamental form of S_{n+p} is denoted by $\sum_\alpha c_\alpha (dz^\alpha)^2$, where the c's are plus or minus one as the case may be, we have

$$\sum_\alpha c_\alpha z^\alpha_{,i} z^\alpha_{,j} = g_{ij}, \qquad z^\alpha_{,i} \equiv \frac{\partial z^\alpha}{\partial x^i}. \tag{1.1}$$

Differentiating covariantly with respect to x^l, we have

$$\sum_\alpha c_\alpha z^\alpha_{,il} z^\alpha_{,j} + \sum_\alpha c_\alpha z^\alpha_{,i} z^\alpha_{,jl} = 0.$$

[1] *Zur Theorie der Transformation quadratischer Differentialausdrücke und der Krümming hoherer Mannigfaltigkeiten*, Math. Annalen, vol. 16 (1880), pp. 129–179.

[2] *Formole fondamentali nella teoria generale delle varietà e della loro curvature*, Rendiconti dei Lincei, ser. 5, vol. 11¹ (1902), pp. 355–362.

[3] R. G., pp. 159–163, 189–192. A reference of this sort is to the author's *Riemannian Geometry*, Princeton, 1926.

[4] Cf. Monatshefte für Mathematik und Physik, vol. 34 (1926), pp. 89–136; vol. 35 (1928), pp. 87–110; vol. 36 (1929), pp. 97–130; also Transactions of the Amer. Math. Soc., vol. 38 (1935), pp. 267–309.

If we subtract this equation from the sum of the two equations obtained from it by interchanging i and l, and j and l respectively, we obtain

$$\sum_{\alpha} c_\alpha z^\alpha_{,l} z^\alpha_{,ij} = 0. \tag{1.2}$$

It follows from these equations that $z^\alpha_{,ij}$ are zero or the components of a vector field of normals to V_n. But from the Ricci identities[5]

$$z^\alpha_{,ijk} - z^\alpha_{,ikj} = z^\alpha_{,h} R^h_{ijk} \tag{1.3}$$

it follows that all of the quantities $z^\alpha_{,ij}$ cannot be zero, since $R_{hijk} \neq 0$ by hypothesis. When i and j take all values 1 to n, there are at most $\frac{1}{2}n(n+1)$ such vectors; we denote by q_1 ($\leqq p$) the least number of independent normals in terms of which the former are linearly expressible. Hence we have[6]

$$z^\alpha_{,ij} = \sum_{\rho} a_{\rho|ij} \lambda^\alpha_\rho \qquad (\rho = 1, \cdots, q_1). \tag{1.4}$$

When $q_1 = p$, the vectors λ^α_ρ can be chosen to be mutually orthogonal unit vectors, since it is assumed that the determinant $g \equiv | g_{ij} |$ is not zero.[7]

In what follows we consider the case $q_1 < p$. From (1.2) we obtain

$$\sum_{\alpha} c_\alpha (z^\alpha_{,l} z^\alpha_{,ijk} + z^\alpha_{,lk} z^\alpha_{,ij}) = 0. \tag{1.5}$$

If λ^α_ρ in (1.4) are mutually orthogonal null-vectors, we have from (1.5) that

$$\sum_{\alpha} c_\alpha z^\alpha_{,l} z^\alpha_{,ijk} = 0. \tag{1.6}$$

Multiplying (1.3) by $c_\alpha z^\alpha_{,l}$ and summing for α, we have in consequence of (1.6) that $R_{hijk} = 0$, contrary to hypothesis. Hence the vectors λ^α_ρ, even if null vectors, are not mutually orthogonal, and consequently there is at least one non-null vector which is linearly expressible in terms of λ^α_ρ; we choose its components so that it is a unit vector, and denote it by η^α_1. Either every vector orthogonal to it and expressible linearly in terms of λ^α_ρ is expressible in terms of mutually orthogonal null-vectors or not. If not, we get another unit vector η^α_2. Continuing this process we have that $z^\alpha_{,ij}$ are expressible in terms of $q_1'(\leqq q_1)$ mutually orthogonal unit vectors η^α_σ and $q_1 - q_1'$ mutually orthogonal null vectors ξ^α_ω, each orthogonal to the vectors η^α_σ. We say that these q_1 vector-fields and all vectors linearly expressible in terms of them constitute the *first normal complex*, N_1. We have

$$\sum_{\alpha} c_\alpha (\eta^\alpha_\sigma)^2 = e_\sigma, \qquad \sum_{\alpha} c_\alpha \eta^\alpha_\sigma \eta^\alpha_\tau = 0 \;\; (\sigma \neq \tau), \qquad \sum_{\alpha} c_\alpha \eta^\alpha_\sigma z^\alpha_{,i} = 0 \tag{1.7}$$

$$(\sigma, \tau = 1, \cdots, q_1'),$$

---

[5] Cf. R. G., p. 30.

[6] Cf. R. G., §56.

[7] R. G., p. 145. The case $q_1 = p$ is treated in §56.

where the e's are plus one or minus one as the case may be. Also we have

$$\sum_\alpha c_\alpha(\xi^\alpha_\omega)^2 = 0, \quad \sum_\alpha c_\alpha \xi^\alpha_\omega \eta^\alpha_\sigma = 0, \quad \sum_\alpha c_\alpha \xi^\alpha_\omega \xi^\alpha_{\omega'} = 0, \quad \sum_\alpha c_\alpha \xi^\alpha_\omega z^\alpha_{,i} = 0 \tag{1.8}$$
$$(\omega, \omega' = q_1' + 1, \cdots, q_1).$$

We replace (1.4) by

$$z^\alpha_{,ij} = \sum_\sigma e_\sigma b_{\sigma/ij}\eta^\alpha_\sigma + \sum_\omega c_{\omega/ij}\xi^\alpha_\omega. \tag{1.9}$$

If the last of (1.7) is differentiated covariantly with respect to x^k, we have in consequence of (1.9), (1.7) and (1.8)

$$\sum_\alpha c_\alpha \eta^\alpha_{\sigma,k} z^\alpha_{,i} + b_{\sigma/ik} = 0,$$

from which and (1.1) it follows that

$$\sum_\alpha c_\alpha(\eta^\alpha_{\sigma,k} + b_{\sigma/hk} g^{hl} z^\alpha_{,l}) z^\alpha_{,i} = 0. \tag{1.10}$$

Also from the last of (1.8) and (1.9) we have

$$\sum_\alpha c_\alpha \xi^\alpha_{\omega,k} z^\alpha_{,i} = 0. \tag{1.11}$$

Hence the quantities $\xi^k_{\omega,k}$ and the expressions in parenthesis in (1.10) for each set of values of ω, k and σ are components of normals to V_n. These normals, as ω, k and σ take on all possible values, are expressible linearly in terms of η^α_σ, ξ^α_ω and $q_2 - q_1(\leqq p)$ other independent normals, say $\xi^\alpha_{\sigma_2}$ for $\sigma_2 = q_1 + 1, \cdots, q_2$.

Since any normal to V_n is expressible linearly in terms of the unit vectors η^α_σ and $p - q_1'$ other mutually orthogonal unit vectors orthogonal to η^α_σ[8], it follows that each vector $\xi^\alpha_{\sigma_2}$ is orthogonal to the vectors η^α_σ. Consequently we have

$$\sum_\alpha c_\alpha \eta^\alpha_\sigma \xi^\alpha_{\sigma_2} = 0, \quad \sum_\alpha c_\alpha \xi^\alpha_{\sigma_2} z^\alpha_{,i} = 0 \quad (\sigma_2 = q_1 + 1, \cdots, q_2). \tag{1.12}$$

In view of the preceding discussion we have

$$\eta^\alpha_{\sigma,k} = -b_{\sigma/jk} g^{jh} z^\alpha_{,h} + \sum_\tau e_\tau \nu_{\tau\sigma/k} \eta^\alpha_\tau + \sum_\omega \mu_{\omega\sigma/k} \xi^\alpha_\omega + \sum_{\sigma_2} \mu_{\sigma_2\sigma/k} \xi^\alpha_{\sigma_2} \tag{1.13}$$

and

$$\xi^\alpha_{\omega,k} = \sum_\tau e_\tau \nu_{\tau\omega/k} \eta^\alpha_\tau + \sum_{\omega'} \mu_{\omega'\omega/k} \xi^\alpha_{\omega'} + \sum_{\sigma_2} \mu_{\sigma_2\omega/k} \xi^\alpha_{\sigma_2} \tag{1.14}$$
$$(\sigma, \tau = 1, \cdots, q_1'; \omega, \omega' = q_1' + 1, \cdots, q_1),$$

the ν's and μ's so defined being components of covariant vectors in V_n.

We say that the vector-fields $\xi^\alpha_{\sigma_2}$, if they exist, span the *second normal complex*, N_2. This normal complex is determined by $z^\alpha_{,ijk}$ obtained from (1.9), and

---

[8] R. G., p. 145.

consequently is not conditioned by the arbitrariness involved in the choice of the vectors η^α_σ and ξ^α_ω.

If we put

$$\sum_\alpha c_\alpha \xi^\alpha_\omega \xi^\alpha_{\sigma_2} = a_{\omega\sigma_2}, \tag{1.15}$$

we have, on differentiating covariantly with respect to x^k the first three of equations (1.8) and making use of (1.12), (1.13) and (1.14),

$$\sum_{\sigma_2} \mu_{\sigma_2\omega/k} a_{\omega\sigma_2} = 0, \tag{1.16}$$

$$\sum_{\sigma_2} \mu_{\sigma_2\sigma/k} a_{\omega\sigma_2} + \nu_{\sigma\omega/k} = 0, \tag{1.17}$$

$$\sum_{\sigma_2} (\mu_{\sigma_2\omega/k} a_{\omega'\sigma_2} + \mu_{\sigma_2\omega'/k} a_{\omega\sigma_2}) = 0. \tag{1.18}$$

If we express the condition of integrability of (1.9) by means of (1.3), and observe that there can be no linear relation between $z^\alpha_{,i}$, η^α_σ, ξ^α_ω and $\xi^\alpha_{\sigma_2}$, we obtain the following four sets of equations of condition, in which a term with carets (^) over two indices stands for this term minus the similar term with these indices interchanged as in (1.19):

$$R_{hijk} = \sum_\sigma e_\sigma(b_{\sigma/hj} b_{\sigma/ik} - b_{\sigma/hk} b_{\sigma/ij}) \equiv \sum_\sigma e_\sigma b_{\sigma/h\hat{j}} b_{\sigma/i\hat{k}}, \tag{1.19}$$

$$b_{\sigma/i\hat{j},\hat{k}} + \sum_\tau e_\tau b_{\tau/i\hat{j}} \nu_{\sigma\tau/\hat{k}} + \sum_\omega c_{\omega/i\hat{j}} \nu_{\sigma\omega/\hat{k}} = 0, \tag{1.20}$$

$$c_{\omega/i\hat{j},\hat{k}} + \sum_\sigma e_\sigma b_{\sigma/i\hat{j}} \mu_{\omega\sigma/\hat{k}} + \sum_{\omega'} c_{\omega'/i\hat{j}} \mu_{\omega\omega'/\hat{k}} = 0, \tag{1.21}$$

$$\sum_\sigma e_\sigma b_{\sigma/i\hat{j}} \mu_{\sigma_2\sigma/\hat{k}} + \sum_\omega c_{\omega/i\hat{j}} \mu_{\sigma_2\omega/\hat{k}} = 0. \tag{1.22}$$

If we differentiate the second of equations (1.12) covariantly with respect to x^k, the resulting equations are reducible by means of (1.9) and (1.12) to

$$\sum_\alpha c_\alpha \xi^\alpha_{\sigma_2,k} z^\alpha_{,i} + \sum_\omega c_{\omega/ik} \sum_\alpha c_\alpha \xi^\alpha_{\sigma_2} \xi^\alpha_\omega = 0$$

from which we have, in consequence of (1.1) and (1.15),

$$\sum_\alpha c_\alpha \Big(\xi^\alpha_{\sigma_2,k} + \sum_\omega c_{\omega/jk} a_{\omega\sigma_2} g^{jh} z^\alpha_{,h}\Big) z^\alpha_{,i} = 0.$$

Consequently the expressions in parentheses are components of a normal vector. Hence we have

$$\begin{aligned} \xi^\alpha_{\sigma_2,k} = -\sum_\omega c_{\omega/jk} a_{\omega\sigma_2} g^{jh} z^\alpha_{,h} + \sum_\sigma e_\sigma \nu_{\sigma\sigma_2/k} \eta^\alpha_\sigma + \sum_\omega \mu_{\omega\sigma_2/k} \xi^\alpha_\omega \\ + \sum_{\tau_2} \mu_{\tau_2\sigma_2/k} \xi^\alpha_{\tau_2} + \sum_{\sigma_3} \mu_{\sigma_3\sigma_2/k} \xi^\alpha_{\sigma_3} \qquad (\sigma_3 = q_2 + 1, \cdots, q_3), \end{aligned} \tag{1.23}$$

where $\xi^\alpha_{\sigma_3}$, if they exist, are the components of normal vector-fields spanning the *third normal complex*, N_3.

From considerations similar to those leading to the first of (1.12) we have that

$$\sum_{\alpha} c_\alpha \eta^\alpha_\sigma \xi^\alpha_{\sigma_3} = 0. \tag{1.24}$$

From (1.7), (1.8), (1.12) and (1.13) we obtain

$$\nu_{\sigma\tau/k} + \nu_{\tau\sigma/k} = 0 \qquad (\sigma, \tau = 1, \cdots, q_1'). \tag{1.25}$$

When we express the condition of integrability of $\eta^\alpha_{\sigma/,k}$, that is $\eta^\alpha_{\sigma,kl} = \eta^\alpha_{\sigma,lk}$, we obtain in consequence of (1.20) and (1.25)

$$\sum_{\omega} \Big[c_{\omega/j\hat{k}} \nu_{\sigma\omega/\hat{i}} + \sum_{\sigma_2} c_{\omega/j\hat{k}} \mu_{\sigma_2\sigma/\hat{i}} a_{\omega\sigma_2} \Big] = 0, \tag{1.26}$$

$$\begin{gathered} \nu_{\tau\sigma/\hat{k},\hat{i}} + g^{hi} b_{\tau/h\hat{k}} b_{\sigma/i\hat{i}} + \sum_{\rho} e_\rho \nu_{\rho\sigma/\hat{k}} \nu_{\tau\rho/\hat{i}} + \sum_{\epsilon} \mu_{\epsilon\sigma/\hat{k}} \nu_{\tau\epsilon/\hat{i}} = 0. \\ (\epsilon = q_1' + 1, \cdots, q_2), \end{gathered} \tag{1.27}$$

$$\begin{gathered} \mu_{\omega\sigma/\hat{k},\hat{i}} + g^{hi} c_{\omega/h\hat{k}} b_{\sigma/i\hat{i}} + \sum_{\tau} e_\tau \nu_{\tau\sigma/\hat{k}} \mu_{\omega\tau/\hat{i}} + \sum_{\epsilon} \mu_{\epsilon\sigma/\hat{k}} \mu_{\omega\epsilon/\hat{i}} = 0 \\ (\epsilon = q_1' + 1, \cdots, q_2), \end{gathered} \tag{1.28}$$

$$\mu_{\sigma_2\sigma/\hat{k},\hat{i}} + \sum_{\tau} e_\tau \nu_{\tau\sigma/\hat{k}} \mu_{\sigma_2\tau/\hat{i}} + \sum_{\epsilon} \mu_{\epsilon\sigma/\hat{k}} \mu_{\sigma_2\epsilon/\hat{i}} = 0. \tag{1.29}$$

$$\sum_{\sigma_2} \mu_{\sigma_2\sigma/\hat{k}} \mu_{\sigma_3\sigma_2/\hat{i}} = 0. \tag{1.30}$$

Proceeding in the same manner with (1.14) we obtain

$$\sum_{\tau} e_\tau \nu_{\tau\omega/\hat{k}} b_{\tau/h\hat{i}} + \sum_{\sigma_2,\omega'} \mu_{\sigma_2\omega/\hat{k}} c_{\omega'/h\hat{i}} a_{\omega'\sigma_2} = 0 \qquad (\omega, \omega' = q_1' + 1, \cdots, q_1), \tag{1.31}$$

$$\nu_{\tau\omega/\hat{k},\hat{i}} + \sum_{\rho} e_\rho \nu_{\rho\omega/\hat{k}} \nu_{\tau\rho/\hat{i}} + \sum_{\omega'} \mu_{\omega'\omega/\hat{k}} \nu_{\tau\omega'/\hat{i}} + \sum_{\sigma_2} \mu_{\sigma_2\omega/\hat{k}} \nu_{\tau\sigma_2/\hat{i}} = 0, \tag{1.32}$$

$$\mu_{\epsilon\omega/\hat{k},\hat{i}} + \sum_{\sigma} e_\sigma \nu_{\sigma\omega/\hat{k}} \mu_{\epsilon\sigma/\hat{i}} + \sum_{\omega'} \mu_{\omega'\omega/\hat{k}} \mu_{\epsilon\omega'/\hat{i}} + \sum_{\sigma_2} \mu_{\sigma_2\omega/\hat{k}} \mu_{\epsilon\sigma_2/\hat{i}} = 0, \tag{1.33}$$

$$\sum_{\sigma_2} \mu_{\sigma_2\omega/\hat{k}} \mu_{\sigma_3\sigma_2/\hat{i}} = 0 \qquad (\epsilon = q_1' + 1, \cdots, q_2). \tag{1.34}$$

2. We consider the case when the first normal complex N_1 consists of a single vector-field. By the argument following equation (1.6) this cannot be a null-vector field provided $R_{hijk} \neq 0$. If it is a unit vector-field $\eta^\alpha_{1,}$, it follows from (1.22) that $b_{1/ij} = \rho_i \mu_{\sigma_2 i/j}$, and then from (1.19) that $R_{hijk} = 0$. Hence we have:

The first normal complex N_1 of a V_n for which $R_{hijk} \neq 0$, immersed in a flat-space of order $n + p$ $(p > 1)$ consists of more than one vector-field.[9]

As a corollary we have:

---

[9] This theorem for the case when the fundamental form is definite has been established by Burstin, Monatshefte, vol. 36 (1929), p. 114.

For a V_n of class 2, when immersed in a flat-space of order $n + 2$, all the vectors normal to V_n are in the first normal complex N_1.

We consider next the case when the first normal complex is spanned by q_1' unit vector-fields and $q_1 - q_1'$ null vector-fields, all mutually orthogonal, and N_2 is vacuous. Differentiating the second of equations (1.8) and making use of (1.13) with $\mu_{\sigma_2\sigma/k} = 0$ and (1.14) with $\mu_{\sigma_2\omega/k} = 0$, we obtain

$$\nu_{\sigma\omega/k} = 0. \tag{2.1}$$

In this case (1.20) and (1.27) reduce to

$$b_{\sigma/i\hat{j},\hat{k}} + \sum_\tau e_\tau b_{\tau/i\hat{j}} \nu_{\sigma\tau/\hat{k}} = 0, \tag{2.2}$$

$$\nu_{\tau\sigma/\hat{k},\hat{l}} + g^{hi} b_{\tau/h\hat{k}} b_{\sigma/i\hat{l}} + \sum_\rho e_\rho \nu_{\rho\sigma/\hat{k}} \nu_{\tau\rho/\hat{l}} = 0. \tag{2.3}$$

But these and (1.19) are the conditions of integrability of

$$z^\alpha_{,ij} = \sum_\sigma e_\sigma b_{\sigma/ij} \eta^\alpha_\sigma, \tag{2.4}$$

$$\frac{\partial \eta^\alpha_\sigma}{\partial x^k} = -b_{\sigma/hk} g^{hj} z^\alpha_{,j} + \sum_\tau e_\tau \nu_{\tau\sigma/k} \eta^\alpha_\tau \quad (\sigma, \tau = 1, \cdots, q_1'). \tag{2.5}$$

[10]

Hence we have

If a V_n is immersed in a flat-space of order $n + p$ and the first normal complex is spanned by q_1' unit vector fields and $q_1 - q_1'$ null vector-fields all mutually orthogonal, and if the second vector complex N_2 is vacuous, then V_n can be immersed in a flat-space of $n + q_1'$ dimensions, and the first normal complex is spanned by q_1' mutually orthogonal unit vector-fields.

Consider now the case when in the preceding theorem the normal complex N_2 is not vacuous, and the number of independent equations in the Pfaffian system

$$\nu_{\tau\omega/k} dx^k = 0, \qquad \mu_{\epsilon\omega/k}\, dx^k = 0 \qquad (\tau = 1, \cdots, q_1';\ \epsilon = q_1' + 1, \cdots, q_2) \tag{2.6}$$

for a given value of ω is r which we call the *rank*. If $r = n - 1$, the equations define a congruence of curves in V_n, such that the components of ξ^α_ω at all points of any one of these curves have the same value, as follows from (1.14), that is, these normals are parallel. If $r < n - 1$, and dx^k and δx^k are two sets of differentials satisfying (2.6), we have from (1.32) and (1.33)

$$\nu_{\tau\omega/\hat{k},\hat{l}}\, dx^k \delta x^l = 0, \qquad \mu_{\epsilon\omega/\hat{k},\hat{l}}\, dx^k \delta x^l = 0.$$

Consequently the Pfaffian system (2.6) is completely integrable, that is it admits a solution

$$\varphi_a(x^1, \cdots, x^n) = c_a \qquad (a = 1, \cdots, n - r), \tag{2.7}$$

where the c's are constants.[11] In this case the normals ξ^α_ω at all points of each V_r defined by (2.7) are parallel. Hence we have

[10] Cf. R. G., p. 190.

[11] Cf. Goursat, *Le Problème de Pfaff*, Hermann, Paris, 1922, p. 267.

If for a given ω the rank of the Pfaffian system (2.6) *is* $r(< n)$, *the system defines in* V_n *a family of sub-varieties* V_r, *one through each point, such that the normals* ξ^α_ω *to* V_n *at all points of a* V_r *are parallel.*

If the vector ξ^α_ω is orthogonal to all the vectors $\xi^\alpha_{\sigma_2}$, we have (2.1) and the Pfaffian system is $\mu_{\epsilon\omega/k}\, dx^k = 0$. If all the vectors ξ^α_ω for $\omega = q_1' + 1, \cdots, q_1$ are orthogonal to a vector $\xi^\alpha_{\sigma_2}$, that is $a_{\omega\sigma_2} = 0$, and the rank of the Pfaffian system

$$\nu_{\sigma\sigma_2/k} dx^k = 0, \qquad \mu_{\epsilon\sigma_2/k} dx^k = 0 \qquad (\epsilon = q_1' + 1, \cdots, q_3)$$

is $r_1(< n)$, there exists a family of varieties V_r, in V_n at all points of each of which the vectors $\xi^\alpha_{\sigma_2}$ are parallel, as follows from (1.23) and the conditions of integrability of these equations.

3. In this section we consider the case when N_2 consists of a single vector-field $\xi^\alpha_{q_2}$ and N_3 is not vacuous. From (1.30) and (1.34) we have

$$\mu_{q_2\sigma/k} = \theta_\sigma \mu_k, \qquad \mu_{q_2\omega/k} = \theta_\omega \mu_k, \tag{3.1}$$

and $\mu_{q_3q_2/k} = \psi_{\sigma_2}\mu_k$. The θ's and ψ's are scalars. Not all the θ's are zero, otherwise N_2 is vacuous, as follows from (1.13) and (1.14). Likewise all the ψ's are not zero, otherwise N_3 is vacuous. From (1.23) it follows that N_3 consists of the single vector-field $\sum_{\sigma_3} \psi_{\sigma_3}\theta^\alpha_{\sigma_3}$, which we denote by $\xi^\alpha_{q_3}$, where $q_3 = q_2 + 1 = q_1 + 2$, and we put

$$\mu_{q_3q_2/k} = \mu_k, \tag{3.2}$$

and use these μ's in (3.1). Thus we have

If N_2 *consists of a single vector-field and* N_3 *is not vacuous, then* N_3 *consists of a single vector-field.*

Consider the quantities $\sum_\alpha e_\alpha \xi^\alpha_\omega \xi^\alpha_{q_2}$. Either all of them are zero, or by replacing the vectors ξ^α_ω by suitable linear combinations of them we have

$$\sum_\alpha c_\alpha \xi^\alpha_{q_1}\xi^\alpha_{q_2} = a_{q_1}, \qquad \sum_\alpha c_\alpha \xi^\alpha_{\bar\omega}\xi^\alpha_{q_2} = 0 \quad (\bar\omega = q_1' + 1, \cdots, q_1 - 1), \tag{3.3}$$

and the components of $\xi^\alpha_{q_1}$ can be chosen so that $a_{q_1} = 1$, if not zero. Now (1.23) becomes

$$\xi^\alpha_{q_2,k} = -c_{q_1/jk}a_{q_1}g^{jh}z^\alpha_{,h} + \sum_\sigma e_\sigma \nu_{\sigma q_2/k}\eta^\alpha_\sigma + \sum_\omega \mu_{\omega q_2/k}\xi^\alpha_\omega + \mu_{q_2q_2/k}\xi^\alpha_{q_2} + \mu_k \xi^\alpha_{q_3}. \tag{3.4}$$

In order to obtain the conditions of integrability of (3.4), we must have an expression for $\xi^\alpha_{q_3,k}$. On differentiating

$$\sum_\alpha c_\alpha \xi^\alpha_{q_3} z^\alpha_{,i} = 0$$

covariantly, we see that, if we put

$$\sum_\alpha c_\alpha \xi^\alpha_\omega \xi^\alpha_{q_3} = b_\omega, \tag{3.5}$$

the quantities $\xi^{\alpha}_{q_3,k}$ are necessarily of the form

$$\xi^{\alpha}_{q_3,k} = -\sum_{\omega} c_{\omega/jk} b_{\omega} g^{jh} z^{\alpha}_{,h} + \sum_{\sigma} e_{\sigma} \nu_{\sigma q_3/k} \eta^{\alpha}_{\sigma} + \sum_{\omega} \mu_{\omega q_3/k} \xi^{\alpha}_{\omega} + \mu_{q_2 q_3/k} \xi^{\alpha}_{q_2} + \mu_{q_3 q_3/k} \xi^{\alpha}_{q_3} + \sum_{\sigma_4} \mu_{\sigma_4 q_3/k} \xi^{\alpha}_{\sigma_4}, \tag{3.6}$$

where the μ's and ν's so defined are components of covariant vectors in V_n, and $\xi^{\alpha}_{\sigma_4}$ for $\sigma_4 = q_3 + 1, \cdots, q_4$ span the normal complex, N_4, if it exists.

Expressing that $\xi^{\alpha}_{q_2,kl} = \xi^{\alpha}_{q_2,lk}$, we obtain

$$c_{q_1/j\dot{k},\dot{l}} a_{q_1} + \sum_{\tau} e_{\tau} \nu_{\tau q_2/\dot{k}} b_{\tau/j\dot{l}} + c_{q_1/j\dot{l}} \mu_{q_2 q_2/\dot{k}} a_{q_1} + \mu_{\dot{k}} \sum_{\omega} c_{\omega/j\dot{l}} b_{\omega} = 0, \tag{3.7}$$

$$\nu_{\sigma q_2/\dot{k},\dot{l}} + a_{q_1} g^{jh} b_{\sigma/h\dot{k}} c_{q_1/j\dot{l}} + \sum_{\tau} e_{\tau} \nu_{\tau q_2/\dot{k}} \nu_{\sigma\tau/\dot{l}} + \sum_{\omega} \mu_{\omega q_2/\dot{k}} \nu_{\sigma\omega/\dot{l}} + \mu_{q_2 q_2/\dot{k}} \nu_{\sigma q_2/\dot{l}} + \mu_{\dot{k}} \nu_{\sigma q_3/\dot{l}} = 0, \tag{3.8}$$

$$\mu_{\omega q_2/\dot{k},\dot{l}} - a_{q_1} c_{q_1/j\dot{k}} c_{\omega/h\dot{l}} g^{jh} + \sum_{\sigma} e_{\sigma} \nu_{\sigma q_2/\dot{k}} \mu_{\omega\sigma/\dot{l}} + \sum_{\omega'} \mu_{\omega' q_2/\dot{k}} \mu_{\omega\omega'/\dot{l}} + \mu_{q_2 q_2/\dot{k}} \mu_{\omega q_2/\dot{l}} + \mu_{\dot{k}} \mu_{\omega q_3/\dot{l}} = 0, \tag{3.9}$$

$$\mu_{q_2 q_2/\dot{k},\dot{l}} + \sum_{\sigma} e_{\sigma} \theta_{\sigma} \nu_{\sigma q_2/\dot{k}} \mu_{\dot{l}} + \sum_{\omega} \mu_{\omega q_2/\dot{k}} \mu_{q_2\omega/\dot{l}} + \mu_{\dot{k}} \mu_{q_2 q_3/\dot{l}} = 0, \tag{3.10}$$

$$\mu_{\dot{k},\dot{l}} + \mu_{q_2 q_2/\dot{k}} \mu_{\dot{l}} + \mu_{\dot{k}} \mu_{q_3 q_3/\dot{l}} = 0, \tag{3.11}$$

$$\mu_{\dot{k}} \mu_{\sigma_4 q_3/\dot{l}} = 0. \tag{3.12}$$

From the last of these equations it follows that in fact N_4 consists of a single vector-field, if at all. In the same way we find on considering the integrability of $\xi^{\alpha}_{q_3,k}$ that N_5 consists of a single vector. In fact, if N_2 consisted of more than one vector-field, and N_3 of one vector-field, we should have found that $N_4, \cdots$ consisted of one vector-field each. Hence we have:

If any normal complex other than the first consists of one vector-field, the same is true of all subsequent complexes until a vacuous one is reached.[12]

We consider the Pfaffian equation

$$\mu_k \, dx^k = 0. \tag{3.13}$$

If dx^k and δx^k are two sets of differentials satisfying this equation, it follows from (3.11) that

$$\mu_{k,l} \, dx^k \delta x^l = 0.$$

Hence equation (3.13) admits an integrating factor.[13] Since μ_k as introduced by (3.1) and (3.2) was determined to within a factor we may in all generality take μ_k as a gradient, thus

$$\mu_k = \frac{\partial \varphi}{\partial x^k}. \tag{3.14}$$

[12] Cf. Burstin, l. c.

[13] Cf. Goursat, l. c., p. 18.

Hence equations (3.11) reduce to

$$\mu_k^i(\mu_{q_3q_3/i} - \mu_{q_2q_2/i}) = 0. \tag{3.15}$$

Differentiating the first of (1.12) and making use of (1.13), (2.1), (2.3) and (2.4), we obtain

$$\nu_{\sigma q_2/k} + a_{q_1}\mu_{q_1\sigma/k} + \theta_\sigma \mu_k \sum_\alpha c_\alpha(\xi^\alpha_{q_2})^2 = 0. \tag{3.16}$$

If $a_{q_1} = 0$ in (3.3), in which case $\xi^\alpha_{q_2}$ is orthogonal to all of the vectors ξ^α_ω, it follows from (1.17) and (3.16) that

$$\nu_{\sigma\omega/k} = 0, \qquad \nu_{\sigma q_2/k} = -\theta_\sigma \mu_k \sum_\alpha c_\alpha(\xi^\alpha_{q_2})^2. \tag{3.17}$$

In this case equations (1.20) and (1.27) reduce to (2.2) and (2.3) respectively, which with (1.19) are the conditions of integrability of (2.4) and (2.15). Hence we have

If a V_n is immersed in a flat space and the first normal complex N_1 is spanned by q_1' unit vector-fields and $q_1 - q_1'$ null vector-fields all mutually orthogonal, and if N_2 consists of a single vector-field orthogonal to the null-vectors of N_1 as it is also to the unit vectors of N_1, and if N_3 is not vacuous, then V_n can be immersed in a flat space of $n + q_1'$ dimensions, and the first vector-complex is spanned by unit vectors.

As a corollary we have

If a V_n is immersed in a flat space and the first normal complex is spanned by q_1 mutually orthogonal unit vector-fields and N_2 consists of one vector-field and N_3 is not vacuous, then V_n can be immersed in a flat space of $n + q_1$ dimensions.

Although under these conditions V_n can be immersed in a space of $n + q_1$ dimensions, we consider the case when it is so immersed that an N_2 and N_3 exist each containing vector-fields $\xi^\alpha_{q_2}$ and $\xi^\alpha_{q_3}$ respectively. If the former is not a null vector-field, the components can be chosen so that

$$\sum_\alpha c_\alpha(\xi^\alpha_{q_2})^2 = e_{q_2}, \tag{3.18}$$

where e_{q_2} is $+1$ or -1 as the case may be. If the equation

$$\sum_\alpha c_\alpha \xi^\alpha_{q_2}\xi^\alpha_{q_3} = 0 \tag{3.19}$$

is not satisfied, and the left-hand member is equal to b, on taking for $\xi^\alpha_{q_3}$ the vector $\xi^\alpha_{q_3} - be_{q_2}\xi^\alpha_{q_2}$, the condition (3.19) is satisfied. Differentiating equation (3.18), it follows from (3.4) and (3.19) that $\mu_{q_2q_2/k} = 0$, in consequence of which and the second of (3.7) we have from (3.4) and (3.14)

$$\xi^\alpha_{q_2,k} = \left(-\sum_\sigma e_\sigma\theta_\sigma\eta^\alpha_\sigma e_{q_2} + \xi^\alpha_{q_3}\right)\frac{\partial\varphi}{\partial x^k}. \tag{3.20}$$

Consequently $\xi^\alpha_{q_2}$ and the expression in parenthesis are functions of φ.

From (3.10) and (3.11) we have

$$\mu_{q_2 q_3/k} = \rho_2 \mu_k, \qquad \mu_{q_3 q_3/k} = \rho_3 \mu_k, \tag{3.21}$$

where ρ_2 and ρ_3 are scalars. In consequence of these results and (3.12) equations (3.6) become in this case

$$\xi^\alpha_{q_3,k} = \sum_\sigma e_\sigma \nu_{\sigma q_3/k} \eta^\alpha_\sigma + (\rho_2 \xi^\alpha_{q_2} + \rho_3 \xi^\alpha_{q_3} + \psi_4 \xi^\alpha_{q_4}) \frac{\partial \varphi}{\partial x^k}.$$

Differentiating the equations $\sum_\alpha c_\alpha \eta^\alpha_\sigma \xi^\alpha_{q_3} = 0$, we find that $\nu_{\sigma q_3/k} = 0$, and consequently $\xi^\alpha_{q_3}$ and the quantities in parenthesis are functions of φ. Hence we have

If a V_n is immersed in a flat space so that the first normal complex is spanned by q_1 mutually orthogonal unit vector fields η^α_σ, and if the second normal complex consists of a single unit vector-field $\xi^\alpha_{q_1+1}$, and the third normal complex is not vacuous and thus consists of a single vector-field $\xi^\alpha_{q_1+2}$, there exists in V_n a family of hyper-surfaces $\varphi = const.$ such that the normals $\xi^\alpha_{q_1+1}$ and $\xi^\alpha_{q_1+2}$ to V_n at all points of a hypersurface are parallel; the same is true of the normal $\sum_\sigma e_\sigma \theta_\sigma \eta^\alpha_\sigma$, where θ_σ is defined by (3.1).

4. We have remarked that when $a_{q_1} \neq 0$ in (3.3) the vectors ξ^α_ω can be chosen so that

$$\sum_\alpha c_\alpha \xi^\alpha_{q_1} \xi^\alpha_{q_2} = 1, \qquad \sum_\alpha c_\alpha \xi^\alpha_{\bar\omega} \xi^\alpha_{q_2} = 0 \quad (\bar\omega = q_1' + 1, \cdots, q_1 - 1). \tag{4.1}$$

If $\sum_\alpha c_\alpha (\xi^\alpha_{q_2})^2 \neq 0$, the components of $\xi^\alpha_{q_2}$ can be chosen so that

$$\sum_\alpha c_\alpha (\xi^\alpha_{q_2})^2 = 1 \tag{4.2}$$

or -1, and the components of $\xi^\alpha_{q_1}$ chosen so that (4.1) holds. If $\sum_\alpha c_\alpha (\xi^\alpha_{q_2})^2 = -1$, and we replace $\xi^\alpha_{q_2}$ by $\xi^\alpha_{q_2} + \xi^\alpha_{q_1}$, the new $\xi^\alpha_{q_2}$ satisfy (4.1) and (4.2). If $\sum_\alpha c_\alpha (\xi^\alpha_{q_2})^2 = 0$, and we replace $\xi^\alpha_{q_2}$ by $\xi^\alpha_{q_2} + \frac{1}{2} \xi^\alpha_{q_1}$, the new vector satisfies (4.1) and (4.2). Hence we may assume in all generality that (4.1) and (4.2) hold.

There is no loss of generality in assuming that

$$\sum_\alpha c_\alpha \xi^\alpha_{q_1} \xi^\alpha_{q_3} = 1, \qquad \sum_\alpha c_\alpha \xi^\alpha_{q_2} \xi^\alpha_{q_3} = 0. \tag{4.3}$$

In fact, if the left hand number of the first equation were zero, on replacing $\xi^\alpha_{q_3}$ by $\xi^\alpha_{q_3} + \xi^\alpha_{q_2}$, we get the first equation in consequence of (4.1); and if the left hand number were not zero, the components $\xi^\alpha_{q_3}$ could be chosen to give the equation this form. After this has been done, if the left-hand number of the second equation were equal to σ ($\neq 0$), and we replace $\xi^\alpha_{q_3}$ by $\xi^\alpha_{q_3} - \sigma \xi^\alpha_{q_1}$, the new vector $\xi^\alpha_{q_3}$ satisfies both of equations (4.3).

If we define β by $\sum_\alpha c_\alpha (\xi^\alpha_{q_3})^2 = \beta$, and we replace $\xi^\alpha_{q_3}$ by

$$\left(1 - \frac{d}{\sqrt{|\beta|}}\right)(\xi^\alpha_{q_2} - \xi^\alpha_{q_1}) + \frac{d}{\sqrt{|\beta|}} \xi^\alpha_{q_3},$$

where α is a constant ($\neq 1$), then for the new vector the left-hand number of this equation is equal to $ed^2 - 1$, where e is $+1$ or -1 as β is positive or negative, and equations (4.3) are satisfied. Hence we may assume that $\xi^\alpha_{q_3}$ has been chosen so that

$$\sum_\alpha c_\alpha(\xi^\alpha_{q_3})^2 = \alpha_3(\neq 0), \tag{4.4}$$

where α_3 is a constant.

We consider the case when [cf. (3.5)]

$$b_{\bar\omega} \equiv \sum_\alpha c_\alpha \xi^\alpha_{\bar\omega} \xi^\alpha_{q_3} = 0 \qquad (\bar\omega = q_1' + 1, \cdots, q_1 - 1). \tag{4.5}$$

From (1.16), (1.18) and (4.1) we have

$$\mu_{q_2\omega/k} = 0. \tag{4.6}$$

From (1.17), (3.3) and (3.1) we obtain

$$\nu_{\sigma q_1/k} = -\theta_\sigma \mu_k, \qquad \nu_{\sigma\bar\omega/k} = 0, \tag{4.7}$$

and from (3.16) and (4.2)

$$\nu_{\sigma q_2/k} + \mu_{q_1\sigma/k} + \theta_\sigma \mu_k = 0. \tag{4.8}$$

If we put for the sake of brevity

$$\mu_{q_2q_2/k} = \nu_k, \tag{4.9}$$

and differentiate equation (4.2), we obtain, on making use of (3.4), (4.1) and (4.3),

$$\mu_{q_1q_2/k} = -\nu_k. \tag{4.10}$$

Also differentiating equations (4.1), we obtain, in consequence of (4.5) and (4.6),

$$\mu_{q_1q_1/k} = -(\nu_k + \mu_k), \qquad \mu_{q_1\bar\omega/k} = 0. \tag{4.11}$$

Equations (1.33) for $\epsilon = \omega = q_1$ reduce by means of (3.14), (4.6), (4.7), (4.8) and (4.11) to

$$\nu_{\dot k,\dot l} = -\mu_{\dot k} \sum_\sigma e_\sigma \theta_\sigma \mu_{q_1\sigma/\dot l} = \mu_{\dot k} \sum_\sigma c_\sigma \theta_\sigma \nu_{\sigma q_2/\dot l}. \tag{4.12}$$

In consequence of these equations and (3.10) it follows from (3.9) for $\omega = q_1$ that

$$\mu_{\dot k}(\mu_{q_1q_3/\dot l} - \nu_{\dot l}) = 0, \tag{4.13}$$

and from (3.10), (4.9) and (4.12)

$$\mu_{\dot k}\mu_{q_2q_3/\dot l} = 0. \tag{4.14}$$

Also (3.15) becomes

$$\mu_{\dot k}(\mu_{q_3q_3/\dot l} - \nu_{\dot l}) = 0. \tag{4.15}$$

In consequence of the above results equations (1.20) and (1.21) for $\omega = q_1$ reduce to

$$(4.16)\qquad \begin{aligned} &b_{\sigma/i\hat{j},\hat{k}} + \sum_{\tau} e_\tau b_{\tau/i\hat{j}} \nu_{\sigma\tau/\hat{k}} - \theta_\sigma c_{q_1/i\hat{j}} \mu_{\hat{k}} = 0, \\ &c_{q_1/i\hat{j},\hat{k}} + \sum_{\sigma} e_\sigma b_{\sigma/i\hat{j}} \mu_{q_1\sigma/\hat{k}} - c_{q_1/i\hat{j}}(\nu_{\hat{k}} + \mu_{\hat{k}}) = 0, \end{aligned}$$

and since $b_{\sigma/ij}$ is symmetric in i and j, we have from (1.22) and (3.1)

$$(4.17)\qquad \sum_{\sigma} e_\sigma \theta_\sigma b_{\sigma/ij} = \gamma \mu_i \mu_j,$$

where γ is a scalar. Equations (1.27) reduce to

$$(4.18)\qquad \nu_{\tau\sigma/\hat{k},\hat{i}} + g^{hi} b_{\tau/h\hat{k}} b_{\sigma/i\hat{i}} + \sum_{\rho} e_\rho \nu_{\rho\sigma/\hat{k}} \nu_{\tau\rho/\hat{i}} - \theta_\tau \mu_{q_1\sigma/\hat{k}} \mu_{\hat{i}} + \theta_\sigma \mu_{\hat{k}} \nu_{\tau q_2/\hat{i}} = 0.$$

Equations (1.28) for $\epsilon = q_1$ reduce to

$$(4.19)\qquad \mu_{q_1\sigma/\hat{k},\hat{i}} + g^{hi} c_{q_1/h\hat{k}} b_{\sigma/i\hat{i}} + \sum_{\tau} e_\tau \nu_{\tau\sigma/\hat{k}} \mu_{q_1\tau/\hat{i}} + \theta_\sigma \nu_{\hat{k}} \mu_{\hat{i}} + (\nu_{\hat{k}} + \mu_{\hat{k}}) \mu_{q_1\sigma/\hat{i}} = 0,$$

and (1.29) in consequence of (1.25), to

$$(4.20)\qquad \mu_{\hat{k}}\Big(\theta_{\sigma,\hat{i}} + \sum_{\tau} e_\tau \theta_\tau \nu_{\sigma\tau/\hat{i}} + \theta_\sigma \nu_{\hat{i}}\Big) = 0.$$

Equations (2.8) become

$$(4.21)\qquad \nu_{\sigma q_2/\hat{k},\hat{i}} + g^{jh} b_{\sigma/h\hat{k}} c_{q_1/j\hat{i}} + \sum_{\tau} e_\tau \nu_{\tau q_2/\hat{k}} \nu_{\sigma\tau/\hat{i}} + \theta_\sigma \nu_{\hat{k}} \mu_{\hat{i}} + \nu_{\hat{k}} \nu_{\sigma q_2/\hat{i}} + \mu_{\hat{k}} \nu_{\sigma q_3/\hat{i}} = 0.$$

We note that the quantities $c_{\bar{\omega}/ij}$, $\mu_{\bar{\omega}\sigma/k}$, $\mu_{\bar{\omega}q_1/k}$ for $\bar{\omega} = q_1' + 1, \cdots, q_1 - 1$ do not appear in (1.19) and in the foregoing equations (4.12) to (4.21). These are the conditions of integrability of the system of equations

$$(4.22)\qquad \begin{aligned} z^\alpha_{,ij} &= \sum_{\sigma} e_\sigma b_{\sigma/ij} \eta^\alpha_\sigma + c_{q_1/ij} \xi^\alpha_{q_1}, \\ \eta^\alpha_{\sigma,k} &= -b_{\sigma/jk} g^{jh} z^\alpha_{,h} + \sum_{\tau} e_\tau \nu_{\tau\sigma/k} \eta^\alpha_\tau + \mu_{q_1\sigma/k} \xi^\alpha_{q_1} + \theta_\sigma \mu_k \xi^\alpha_{q_2}, \\ \xi^\alpha_{q_1,k} &= -\sum_{\sigma} e_\sigma \theta_\sigma \eta^\alpha_\sigma \mu_k - \xi^\alpha_{q_1}(\mu_k + \nu_k), \\ \xi^\alpha_{q_2,k} &= -c_{q_1/jk} g^{jh} z^\alpha_{,h} + \sum_{\sigma} e_\sigma \nu_{\sigma q_2/k} \eta^\alpha_\sigma - \nu_k(\xi^\alpha_{q_1} - \xi^\alpha_{q_2}) + \mu_k \xi^\alpha_{q_3}, \end{aligned}$$

where, in consequence of (3.6), (3.12), (4.3) and (4.5)

$$(4.23)\qquad \begin{aligned} \xi^\alpha_{q_3,k} = -c_{q_1/jk} g^{jh} z^\alpha_{,h} &+ \sum_{\sigma} e_\sigma \nu_{\sigma q_3/k} \eta^\alpha_\sigma + \mu_{q_1 q_3/k} \xi^\alpha_{q_1} + \mu_{q_2 q_3/k} \xi^\alpha_{q_2} \\ &+ \mu_{q_3 q_3/k} \xi^\alpha_{q_3} + \psi_4 \mu_k \xi^\alpha_{q_4} \end{aligned}$$

ψ_4 being a scalar. Consequently V_n can be immersed so that there is only one null-vector, $\xi^\alpha_{q_1}$, in the first normal complex.

Proceeding in a manner similar to that leading to (4.3) and (4.4), we show that we can choose $\xi^\alpha_{q_4}$ so that

$$(4.24)\quad \sum_\alpha c_\alpha \xi^\alpha_{q_1}\xi^\alpha_{q_4} = 1,\quad \sum_\alpha c_\alpha \xi^\alpha_{q_2}\xi^\alpha_{q_4} = \sum_\alpha c_\alpha \xi^\alpha_{q_3}\xi^\alpha_{q_4} = 0,\quad \sum_\alpha c_\alpha(\xi^\alpha_{q_4})^2 = \alpha_4(\neq 0).$$

Differentiating the equation (1.24), we obtain in consequence of the above results

$$(4.25)\qquad \nu_{\sigma q_3/k} = -\mu_{q_1\sigma/k} = \nu_{\sigma q_2/k} + \theta_\sigma \mu_k.$$

It is readily shown that equations (4.19) and (4.21) are consistent with this result and (4.20).

Differentiating equations (4.3) and (4.4), we obtain in consequence of the preceding results

$$(4.26)\qquad \begin{aligned} &\mu_{q_2q_3/k} + \mu_{q_3q_3/k} + \mu_k(\psi_4 - 1) - \nu_k = 0,\\ &\mu_{q_2q_3/k} + \mu_{q_1q_3/k} + \mu_k\alpha_3 - \nu_k = 0,\\ &\mu_{q_1q_3/k} + \alpha_3\mu_{q_3q_3/k} = 0. \end{aligned}$$

We consider first the case when

$$(4.27)\qquad 1 + \alpha_3 \neq 0,$$

then from (4.26) we have

$$(4.28)\qquad \begin{aligned} \mu_{q_1q_3/k} &= \alpha_3\mu_k\left(\frac{\psi_4}{1+\alpha_3} - 1\right),\\ \mu_{q_2q_3/k} &= \nu_k - \frac{\alpha_3\psi_4}{1+\alpha_3}\mu_k,\\ \mu_{q_3q_3/k} &= \mu_k\left(1 - \frac{\psi_4}{1+\alpha_3}\right). \end{aligned}$$

In order that these expressions be consistent with (4.13), (4.14) and (4.15), we must have

$$(4.29)\qquad \nu_k = \rho\mu_k,$$

where ρ is a scalar, from which and (4.22) we have

$$(4.30)\qquad \xi^\alpha_{q_1,k} = -(\lambda^\alpha + \xi^\alpha_{q_1})\frac{\partial\varphi}{\partial x^k},$$

where we have put

$$(4.31)\qquad \lambda^\alpha = \sum_\sigma e_\sigma\theta_\sigma\eta^\alpha_\sigma + \rho\xi^\alpha_{q_1}.$$

From (4.30) and (3.14) it follows that $\xi^\alpha_{q_1}$ and λ_α are functions of φ. Hence we have

The normals $\xi^\alpha_{q_1}$ to V_n at points of a hypersurface φ = const. are parallel, as are also the normals λ^α.

From (4.20) and (4.29) we have

$$\theta_{\sigma,k} + \sum_{\tau} e_\tau \theta_\tau \nu_{\sigma\tau/k} = \bar{\theta}_\sigma \mu_k, \tag{4.32}$$

and from (4.29) and (4.12)

$$\rho_{,k} + \sum_{\sigma} e_\sigma \theta_\sigma \mu_{q_1\sigma/k} = \bar{\rho}\mu_k, \tag{4.33}$$

where $\bar{\theta}_\sigma$ and $\bar{\rho}$ are scalars. In consequence of these results and (4.17), we have from (4.31)

$$\lambda^\alpha_{,k} = -\Big[\gamma\mu_j g^{jh} z^\alpha_{,h} + \rho\lambda^\alpha - \sum_\sigma e_\sigma \bar{\theta}_\sigma \eta^\alpha_\sigma + (\rho - \bar{\rho})\xi^\alpha_{q_1} - \sum_\sigma e_\sigma \theta^2_\sigma \xi^\alpha_{q_2}\Big]\mu_k. \tag{4.34}$$

This result confirms the statement concerning the normal λ^α in the above theorem, but shows also that at points of a hypersurface $\varphi = \text{const.}$ the vectors whose components are given by the quantity in parenthesis are parallel; only when $\gamma = 0$ is the vector normal to V_n.

Also we have from (4.22) and (4.23)

$$(\xi^\alpha_{q_3} - \xi^\alpha_{q_2})_{,k} = (\lambda^\alpha - \alpha_3\xi^\alpha_{q_1} + \bar{\lambda}^\alpha)\mu_k, \qquad \bar{\lambda}^\alpha \equiv \frac{\psi_4}{1+\alpha_3}[\alpha_3(\xi^\alpha_{q_1} - \xi^\alpha_{q_2}) - \xi^\alpha_{q_3} + (1+\alpha_3)\xi^\alpha_{q_4}],$$

and consequently we have

The normals $\xi^\alpha_{q_3} - \xi^\alpha_{q_2}$ to V_n at points of a hypersurface $\varphi = \text{const.}$ are parallel, as are also the normals $\bar{\lambda}^\alpha$.

5. When $\alpha_3 = -1$, we have from (4.26), (4.13), (4.14) and (4.15)

$$\psi_4 = 0, \qquad \mu_{q_1q_2/k} = \mu_{q_3q_3/k} = \nu_k - \rho\mu_k, \qquad \mu_{q_2q_3/k} = \mu_k(1+\rho), \tag{5.1}$$

where ρ is a scalar. From the first of these equations it follows that N_4 is vacuous, and from these equations, (4.23) and (4.25) we have

$$\xi^\alpha_{q_3,k} = -c_{q_1/jk} g^{jh} z^\alpha_{,h} - \sum_\sigma e_\sigma \mu_{q_1\sigma/k}\eta^\alpha_\sigma + (\nu_k - \rho\mu_k)(\xi^\alpha_{q_1} + \xi^\alpha_{q_3}) + (1+\rho)\mu_k\xi^\alpha_{q_2}. \tag{5.2}$$

Expressing the condition of integrability of these equations, we obtain the second of equations (4.16), (4.19) and

$$\nu_{[k,i]} = \sum_\sigma e_\sigma \theta_\sigma \mu_{q_1\sigma/[k}\mu_{i]} = \mu_{[k}\rho_{,i]}. \tag{5.3}$$

If we put $\nu_k = \rho\mu_k + \omega_k$, it follows from (5.3) that ω_k is a gradient $\partial\omega/\partial x^k$, and consequently we have

$$\nu_k = \rho\mu_k + \omega_{,k}. \tag{5.4}$$

If $\omega = 0$, we have

$$(\xi^\alpha_{q_3} - \xi^\alpha_{q_2})_{,k} = (\lambda^\alpha + \xi^\alpha_{q_2} - \xi^\alpha_{q_3})\mu_k,$$

where λ^α is given by (4.3), so that the last theorem of §4 applies to this case also.

When $\omega_{,k} \neq 0$, we have from (5.2) and (4.22)

$$(\xi^\alpha_{q_3} - \xi^\alpha_{q_2} + \xi^\alpha_{q_1})_{,k} = (\omega_{,k} - \mu_k)(\xi^\alpha_{q_3} - \xi^\alpha_{q_2} + \xi^\alpha_{q_1}),$$

from which and (3.14) we have

$$\xi^\alpha_{q_3} - \xi^\alpha_{q_2} + \xi^\alpha_{q_1} = a^\alpha e^{\omega-\varphi}, \tag{5.5}$$

where the a^α are constants. In consequence of (4.1), (4.2), (4.3) and (4.4) with $\alpha_3 = -1$, we have

$$\sum_\alpha e_\alpha (a^\alpha)^2 = 0.$$

Hence we have

The null-vector a^α of constant components, defined by (5.5), *is normal to V_n.*

The results of this section and the preceding one are based on the assumption (4.5). If $b_{\bar\omega} \neq 0$, a vector $\xi^\alpha_{q_1-1}$ can be chosen so that

$$\sum_\alpha c_\alpha \xi^\alpha_{q_1-1} \xi^\alpha_{q_3} = 1, \qquad \sum_\alpha c_\alpha \xi^\alpha_{\bar\omega} \xi^\alpha_{q_3} = 0 \quad (\bar\omega = q_1' + 1, \cdots, q_1 - 2).$$

Clearly this will lead to further results which may be obtained by processes similar to those which have just been used.

PRINCETON UNIVERSITY.

Annals of Mathematics
Vol. 38, No. 4, October, 1937

ANALYTIC COORDINATE SYSTEMS AND ARCS IN A MANIFOLD[1]

By Hassler Whitney

(Received March 4, 1937)

1. **Introduction.** In a differentiable, or analytic, manifold,[2] each point is in a "coordinate system". Two overlapping systems are related by a differentiable, or analytic, transformation, with non-vanishing Jacobian. An important question is: are there "large" coordinate systems, containing say two given points, or a given arc, etc. In the analytic case, does there even exist an analytic arc joining given points? We answer some questions of this sort here. *We always suppose the given manifold may be imbedded in a Euclidean space* E_μ. This is no restriction in the differentiable case; see DM Theorem 1. The question is unsolved in the analytic case; see the Problem stated below.

We shall speak of manifolds of class C^r as in DM, and shall use $r = \omega$ for the analytic case; write $\infty < \omega$. By a *coordinate system* (of class C^r) in a C^r-m-manifold M we shall mean any regular (1-1) C^r-map (see DM, §§2, 3) of a "fundamental domain" Q of Euclidean m-space into M. For Q we shall take a spherical region or m-cube, or the region (for some k)

$$\sum_{i=1}^{k} x_i^2 < 1, \qquad -1 < x_j < 1 \qquad (j = k + 1, \cdots, m). \tag{1}$$

We begin by showing that a finite set of points in M may be joined by a differentiable arc. Next it is proved that any submanifold N of M satisfying certain conditions may be imbedded in a family of submanifolds, which fill out an open set in M. If N is an arc, or an n-cube etc., the conditions are always satisfied; the family of submanifolds then defines a coordinate system containing N. Now if two or more points are given in the analytic manifold M (in E_μ), we first join them by a differentiable arc A, and then find an analytic coordinate system R containing A. In R we may then easily find an analytic arc (in fact, an analytic closed curve) joining them. We end the main part of the paper by showing that an analytic closed curve in an analytic manifold M which bounds a differentiable surface element, also bounds an analytic surface element.

In Appendix I we give a simple proof of the imbedding theorem for closed

---

[1] Presented to the Am. Math. Soc., March, 1937.

[2] See for instance DM. We shall refer to the following papers by the author.

AE: *Analytic extensions* $\cdots$, Trans. Am. Math. Soc., vol. 36 (1934), pp. 63–89.

DM: *Differentiable manifolds*, Annals of Math., vol. 37 (1936), pp. 645–680.

IM: *The imbedding of manifolds* $\cdots$, Annals of Math., vol. 37 (1936), pp. 865–878.

FD: *Functions differentiable on the boundaries of regions*, Annals of Math., vol. 35 (1934), pp. 482–485.

manifolds of class C^r, $r \geqq 2$; a theorem on the projecting of a manifold in E_μ into a subspace $E_{\mu-1}$ is used. In Appendix II we clarify the proof of DM Lemma 16.

The simplest possible non-trivial analytic manifold is one of dimension 1, with just two overlapping coordinate systems. For it, the imbedding problem is equivalent to the following problem in function theory:

PROBLEM. Let $\phi(x)$ be any (real) analytic function for $-1 \leqq x \leqq 1$, with

$$\phi(-1) = -1, \qquad \phi(1) = 1, \qquad \phi'(x) > 0 \quad (-1 \leqq x \leqq 1).$$

Does there exist a function $f(x)$, analytic for $-2 < x < 1$, with $f'(x) > 0$ there, such that $g(x) = f(\phi(x))$, analytic for $-1 < x < 1$, may be extended analytically through $1 \leqq x < 2$, so that $g'(x) > 0$ there?

Much use will be made of the following lemma. See DM §6 for notations.

LEMMA 1. *Let M and N be C^r-m- and C^r-n-manifolds respectively,*[3] $1 < r \leqq \omega$, *and let ϕ be a C^t-homeomorphism of N into (a subset of) M, t finite,* $1 \leqq t < r$. *Then for any positive continuous function $\eta(p)$ in N there is a C^r-homeomorphism of N into M approximating (ϕ, N, t, η).*

Generally N will be an open subset of Euclidean space. For $r = \omega$, the lemma is a consequence of DM, Lemma 22; for any sufficiently close approximation to a C^1-homeomorphism is another one. For general r, we may either apply the lemma to C^r-homeomorphic analytic manifolds, or note that DM Lemma 22 holds with "analytic" replaced by "C^r".

COROLLARY. *Any coordinate system of class C^s in the C^r-m-manifold M is contained in one of class C^r, if $1 \leqq s < r$.*

For a generalization of this theorem, see Theorem 2.

2. Differentiable arcs and simple closed curves in M.

We prove here

LEMMA 2. *Theorem 3 is true if $r = 1$.*

Suppose we have found such an arc A' from p_0 as far as p_i $(0 \leqq i < k)$. If $m = 2$, let K be a differentiable simple closed curve running along close beside A' on one side, and returning on the other. (This is easily constructed.) As M is connected, there is a sequence of coordinate systems $U_1, \cdots, U_l$, such that U_1 contains p_i, U_l contains p_{i+1}, $U_j \cdot U_{j+1} \neq 0$, but no other U_j and U_h have common points. Set $q_0 = p_i$, $q_l = p_{i+1}$, and let q_j be a point of $U_j \cdot U_{j+1}$. Join q_j to q_{j+1} in U_j; this gives a path from p_i to p_{i+1}, from which we may pick out an arc B, differentiable except at a finite number of corners. We may make it start from p_i in a given direction. If $m = 2$ and B cuts A', then there is a first point q' and a last point q'' at which it cuts K; we replace the interior part $q'q''$ of B by one of the arcs of K. If $m \geqq 3$ and B cuts A', we may replace small pieces of B cutting A' by differentiable pieces avoiding A'. We now have an arc A^* from p_0 to p_{i+1}. At each corner, we may replace a small piece of A^* by an arc of a circle (using a coordinate system containing the corner); we find thus a differentiable arc from p_0 to p_{i+1}, and finally to p_k. If we want a simple

[3] We recall the hypothesis that the manifolds may be imbedded in Euclidean space.

closed curve, we take another differentiable arc running along beside the first, and join their ends by arcs of circles.

3. **Manifolds in regular position in M.** Let N be a C^s-n-manifold in the C^r-m-manifold M. We say N is in *regular position*[4] in M if the following condition is satisfied. If M is imbedded in some E_μ (DM Theorem 1), then there exist $m - n$ continuous vector functions $v_i(p)$ in N which are independent for each p, and each of which is tangent to M and normal to N. We may clearly take them orthogonal to each other if we wish.

The definition is independent of the choice of the imbedding. For suppose M is also imbedded in E'_ν, and N is regular position in M for M in E_μ. Staying in E_μ, there is a positive continuous function $\xi(p)$ in M' such that the vectors $\xi(p)v_i(p)$ project, for each p, into differentiable curves $c_i(p)$ in M (DM Lemma 23). Going over to E_μ, these curves define independent vectors $v'_i(p)$, the $(m - n)$-plane of which is independent of the tangent plane to N, for each p; for the maps of M into E_μ and into E'_ν are regular. Projecting each such plane onto the plane normal to N and tangent to M in E'_ν gives the required vector functions in E'_ν.

THEOREM 1. *Let N be a C^r-n-manifold in regular position in the C^r-m-manifold*[3] *$M (1 \leqq r \leqq \omega)$. Then there is an $(m - n)$-parameter family of manifolds $N(c_1, \cdots, c_{m-n})$ $(\mid c_i \mid < 1)$ C^r-homeomorphic with N, which fill out a neighborhood of N in M in a (1-1) way, and such that $N(0, \cdots, 0) = N$.*

REMARKS. We shall define the function $Z(p; c_1, \cdots, c_{m-n})$ of class C^r in the product space of N with the open $(m - n)$-cube $\mid c_i \mid < 1$, so that as p varies, we run over $N(c_1, \cdots, c_{m-n})$, and as $c_1, \cdots, c_{m-n}$ vary, we run over an "$(m - n)$-cube" $Q(p)$ in M. If M is of class C^s, $s > r$, we may make the $N(c_1, \cdots, c_{m-n})$ of class C^s for $(c_1, \cdots, c_{m-n}) \neq (0, \cdots, 0)$ (see IM, Theorem I); or we may make all the $N(c_1, \cdots, c_{m-n})$ of class C^s, and have N lie in the open set which they fill out (compare Theorem 2). If r is finite, we may make the $N(c_1, \cdots, c_{m-n})$ approximate to N as in IM, Theorem I, (3).

To prove the theorem, we suppose M is in E_μ, and replace the $v_i(p)$ defined above by independent vector functions $w_i(p)$ as in IM, §4.[5] The $w_i(p)$ are analytic in an open set in E_μ containing N, and hence are of class C^r in N (DM, Lemma 3). Then if $P^*(p)$ is the $(m - n)$-plane through p determined by the $w_i(p)$, the function $P^*(p)$ of p is of class C^r in N (see DM, §24). From the proof of DM Lemma 21 it is clear that for some positive continuous function $\xi(p)$ in N, the points

$$q = p + \sum \alpha_i w_i(p) \qquad (\mid \alpha_i \mid < \xi(p))$$

fill out a C^r-m-manifold N^* containing N. Moreover, if the $w_i(p)$ approximate to the $v_i(p)$ closely enough and $\xi(p)$ is small enough, then N^* is nearly tangent

[4] Compare IM, §3.

[5] Extend the $v_i(p)$ throughout a neighborhood of N in E_μ, and approximate to them with analytic functions.

to M at each p in N, and projects into M(DM Lemma 23), giving a C^r-m-manifold N' in M (i.e. an open subset of M). Extend $\xi(p)$ throughout a neighborhood of N in E_μ, and approximate to $\frac{1}{2}\xi(p)$ by an analytic function $\zeta(p)$ (DM Lemma 9); then $\zeta(p)$, considered in N alone, is of class C^r. Let $Z(p; c_1, \cdots, c_{m-n})$ be that point of N' corresponding to the point

$$p + \sum c_i \zeta(p) w_i(p)$$

in N^*. Z clearly has the required properties.

4. **Coordinate neighborhoods containing given arcs, etc.** Theorem 2 will follow from the last theorem and the following lemma. We shall call a manifold with boundary a manifold also.

LEMMA 3. *Any manifold N in M is in regular position in M if it is C^1-homeomorphic with an open or closed fundamental domain.*

This will be proved in a paper on "sphere-spaces". Because of its importance, we shall give the proof for the case that N is an open or closed arc. Suppose $v_1(p), \cdots, v_{k-1}(p)$ (or no vector functions) have been defined properly on N; we shall define $v_k(p)$. Taking M in E_μ, let $P(p)$ be, for each p on N, the $(m-k)$-plane through p tangent to M and orthogonal to N and to $v_1(p), \cdots, v_{k-1}(p)$; $P(p)$ varies continuously with p. Cut N into subarcs $A_1, A_2, \cdots$ so that no two planes $P(p)$ and $P(q)$ are perpendicular for p and q on the same A_i, and so that each $A_i (i > 1)$ has exactly one point in common with the preceding arcs. Suppose $v_k(p)$ has been defined on $A' = A_1 + \cdots + A_{i-1}$; we shall define it on A_i. Say $A_i = p_0 q_0$, p_0 in A'. Corresponding to p_0 and $P_0 = P(p_0)$, define the transformation $T_{p,P}$ as in DM, §19. For each p in A_i, let $v_k(p)$ be the vector into which $v_k(p_0)$ is carried by $T_{p,P(p)}$. Proceed in this manner until all the $v_k(p)$ are defined throughout N.

REMARK. It is not hard to show that a simple closed curve in M is in regular position if and only if it "preserves orientation" in M.

THEOREM 2. *Any C^s-n-manifold N in the C^r-m-manifold M $(1 \leqq s \leqq r \leqq \omega)$ which is C^s-homeomorphic with an open or closed fundamental domain is contained in a coordinate system of class C^r. If $s = r$, we may make N correspond to the points $x_{n+1} = \cdots = x_m = 0$; in any case, N will approximate to this plane.*

We may suppose N is open; if it is closed, we may extend the map[6] of the closed fundamental domain Q into N to a map of a slightly larger region Q', and apply the proof to Q'. If the theorem with $s = r$ is proved, we may prove the first part of the theorem as follows. Applying the theorem with r replaced by s gives a coordinate system of class C^s containing N; Lemma 1 (applied to the fundamental domain defining this coordinate system) replaces it by one of class C^r.

Suppose then that Q is open and $s = r$. By Theorem 1, Remark, there is a C^r-homeomorphism $Z(p; x_{n+1}, \cdots, x_m)$ of the product S of Q and an open

[6] Supposing that M is in E, we apply the theorem of FD to each coordinate of the given map. The resulting map of Q' is then projected onto M.

$(m - n)$-cube into M. As Q is a fundamental domain, so is S; the map of S is the required coordinate system.

5. Analytic arcs, etc. We can now prove

THEOREM 3.[7] *Let $p_0, p_1, \cdots, p_k$ be points of the connected C^r-m-manifold*[8] *M $(1 \leqq r \leqq \omega, m \geqq 2)$. Then there is an arc, in fact a simple closed curve, of class C^r, containing them in the given order. We may make the arc (or curve) have given directions at the given points.*

REMARK. If there is given a differentiable arc or simple closed curve through the points, it is clear from the proof that we may make the new arc or curve approximate to the given one as closely as we please, both in position and in direction.

We first find an arc A of class C^1 by Lemma 2. By Theorem 2, there is a coordinate system $(x_1, \cdots, x_m)$ of class C^r containing A[8]. In this system, A will approximate to the x_1-axis; in fact, by choosing a good enough approximation in the proof of Theorem 2, we may make the p_i lie as close to the x_1-axis as we please. Say $p_i = (x_1^i, \cdots, x_m^i)$. Let $f_j(x)$ be a polynomial (in one variable) such that

$$f_j(x_1^i) = x_j^i \qquad (i = 0, \cdots, k; j = 2, \cdots, m). \tag{2}$$

The set of points $(x_1, \cdots, x_m)$ satisfying

$$x_j = f_j(x_1) \qquad (j = 2, \cdots, m)$$

is an arc A^* through the points $(x_1^i, \cdots, x_m^i)$. If they are close enough to the x_1-axis, A^* will lie in the coordinate system. A^* is analytic in the coordinate system, and hence of class C^r in M. If we wish A^* to have given directions at the points, we assign the proper derivatives to the polynomials $f_j(x)$.

Suppose a closed curve is desired. We find an arc A^* as above. By Theorem 2, we may suppose it is the x_1-axis in a coordinate system. Join the ends of a corresponding arc $x_2 = \epsilon, x_3 = \cdots = x_m = 0$ to the ends of this one by semicircles, giving the closed curve C. By Lemma 1 there is an approximating closed curve C' of class C^r. The part A' of C' corresponding to A^* is given by functions $x_j = g_j(x_1)$ $(j = 2, \cdots, m)$; set $\bar{x}_j^i = g_j(x_1^i)$, the x_1^i determining the points p_i. Using the $\bar{x}_j^i$, find polynomials $\bar{f}_j(x)$ as before. The transformation of m-space

$$x_1' = x_1, \qquad x_j' = x_j - \bar{f}_j(x_1) \qquad (j = 2, \cdots, m) \tag{3}$$

carries C' into the required closed curve C^*.

6. Analytic cells with analytic boundaries. Our object is to prove Theorem 4. For the case $r = \omega$, we shall need a lemma from potential theory.

LEMMA 4. *Let $f(p)$ be an analytic function on the unit circle A in the plane.*

[7] This Theorem has recently been obtained under weaker hypotheses by T. Y. Thomas, Annals of Math., vol. 38 (1937), pp. 120.

[8] The existence of the required arc is now evident if $k = 1$.

Then the harmonic function $F(p)$ in the interior taking on the boundary values $f(p)$ is analytic in the closed region. There is a constant $K > 0$ such that if $f(p)$ and its first and second derivatives along A are bounded in absolute value by η, then $F(p)$ and its first partial derivatives are bounded in absolute value by $K\eta$.

The first part is well-known.[9] We shall prove the second part with K chosen as follows: Let $u(\theta)$, $|\theta| \leqq \pi$, denote any differentiable function defined on A, measuring θ from the point $(x, y) = (0, -1)$; let $u'(\theta)$, $u''(\theta)$ denote its first and second derivatives with respect to θ. Then there is a $K \geqq 4$ such that

$$|u'(\theta)| \leqq \zeta \text{ (all } \theta) \text{ implies } \left|\frac{u(\theta) - u(0)}{y + 1}\right| \leqq \tfrac{1}{2}K\zeta \qquad (\theta \geqq \tfrac{1}{3}\pi). \tag{4}$$

Now take any $f(\theta)$, with $|f(\theta)|$, $|f'(\theta)|$, $|f''(\theta)| \leqq \eta$. We shall show first that if $F(x, y)$ is the harmonic function in A with the boundary values $f(\theta)$, then

$$\left|\frac{\partial}{\partial y} F(0, -1)\right| \leqq K\eta. \tag{5}$$

Define $G(x, y)$ by

$$G(x, y) = F(x, y) - F(0, -1) - x\frac{\partial}{\partial x} F(0, -1); \tag{6}$$

Then G is harmonic within A, and

$$G(0, -1) = \frac{\partial}{\partial x} G(0, -1) = 0.$$

Let $g(\theta)$ denote the boundary values of G; then

$$g(\theta) = f(\theta) - f(0) - f'(0) \sin \theta, \qquad g'(0) = 0.$$

Using a law of the mean twice (see for instance Osgood, *Advanced Calculus*, p. 209) with $0 < \theta < \pi$ gives, for some α and β,

$$\frac{g(\theta)}{y + 1} = -\frac{g(\theta) - g(0)}{\cos\theta - \cos 0} = -\frac{g'(\alpha)}{-\sin\alpha} \qquad (0 < \alpha < \theta)$$

$$= \frac{g'(\alpha) - g'(0)}{\sin\alpha - \sin 0} = \frac{g''(\beta)}{\cos\beta} \qquad (0 < \beta < \alpha).$$

A similar relation holds with $-\pi < \theta < 0$. Clearly $|g''(\theta)| \leqq 2\eta$; hence

$$\left|\frac{g(\theta)}{y + 1}\right| \leqq \frac{2\eta}{\cos\frac{1}{3}\pi} = 4\eta \qquad (0 < |\theta| \leqq \tfrac{1}{3}\pi). \tag{7}$$

Using the definition of K with this inequality gives

$$\left|\frac{g(\theta)}{y + 1}\right| \leqq K\eta \qquad (\theta \neq 0). \tag{8}$$

[9] See for instance W. F. Osgood, *Funktionentheorie*, 5th. ed. (1928), p. 703, 3. Satz. It must be pointed out that $F(p)$ may be defined so as to be analytic throughout a region containing A, not merely through a neighborhood of any arc of A.

Set

$$H_{\pm}(x, y) = G(x, y) \pm K\eta(y + 1); \tag{9}$$

Then $H_{\pm}$ is harmonic within A. If $h_{\pm}(\theta)$ denotes its boundary values, then (8) gives

$$h_{+}(\theta) \geqq 0, \qquad h_{-}(\theta) \leqq 0 \qquad \text{(all } \theta\text{)}.$$

Therefore the same inequalities hold for H_{+} and H_{-} within A. As they vanish at $(0, -1)$, we have

$$\frac{\partial}{\partial y} H_{+}(0, -1) \geqq 0, \qquad \frac{\partial}{\partial y} H_{-}(0, -1) \leqq 0. \tag{10}$$

Hence (5) holds for G, and therefore also for F. By symmetry, the normal derivative of F at each point of A is bounded by $K\eta$; hence the same is true of any partial derivative of F on A.

Now suppose there were a point p within A at which some partial derivative, say $\partial F/\partial y$, were not bounded by $K\eta$. As $\partial F/\partial y$ is harmonic within A and continuous on the boundary, we have a contradiction with the above inequality and the maximum-minimum theorem.

THEOREM 4. *Let C be a simple closed curve of class C^r in the C^r-m-manifold*[3] *$M(1 \leqq r \leqq \omega)$. If C C^1-bounds*[10] *the 2-cell D, then it C^r-bounds a 2-cell D^*.*

Let A be the unit circle in the plane, with interior B, and let ϕ be the given C^1-homeomorphism of $A + B$ into D. Suppose first that r is finite or $r = \infty$. We may suppose ϕ defined over a slightly larger region B' (compare footnote 6). Let ψ be a C^r-map of B' into M approximating $(\phi, B', 1, \eta)$ for some $\eta > 0$ (Lemma 1). M being in E_μ, ϕ and ψ each have μ coordinates ϕ_i and ψ_i ; consider the functions

$$f_i(p) = \phi_i(p) - \psi_i(p) \qquad (i = 1, \cdots, \mu) \tag{11}$$

on A. Extend the $f_i(p)$ through a neighborhood R of A by letting them be constant on the radii; then they are clearly of class C^r in R (see the proof of DM, Lemma 4). This defines the $f_i(p)$ and their partial derivatives of order $\leqq r$ on A; the $f_i(p)$ are now of class C^r on A, as a subset of the plane (see AE §3). Extend them so as to be of class C^r throughout B', by AE Lemma 2. It is clear from the proof of the lemma referred to that there is a constant K such that if the $f_i(p)$ and their first derivatives are bounded by $\eta > 0$, then the same is true of the extensions $F_i(p)$ of the $f_i(p)$, with η replaced by $K\eta$. Set

$$\Phi_i(p) = \psi_i(p) + F_i(p) \text{ in } A + B.$$

Then $\Phi_i(p)$ is of class C^r in B', and equals $\phi_i(p)$ in A. This set of functions maps B' into E_μ, A going into C; if η and hence $K\eta$ is small enough, this projects (see DM, Lemma 23) into the required cell in M.

---

[10] See DM, §13. Except for the case that $r = \omega$, similar theorems are easily proved in n dimensions. Note that, by the definition in DM, the function ϕ below coincides in A with the function defining C; hence ϕ, considered in A alone, is of class C^r.

Suppose now that $r = \omega$. We first replace $\phi(p)$ by a map of class C^2, by the above proof; call this ϕ again. Define ψ and the $f_i(p)$ again; the $f_i(p)$ are now analytic on A. Extend them through some B' containing $A + B$ by Lemma 4. Making the $f_i(p)$ and their first and second partial derivatives $< \eta$ in absolute value makes the $F_i(p)$ and their first partial derivatives $< K\eta$ in absolute value. Again map B' into E_μ, using the functions $\Phi_i(p)$, and project onto M.[11]

Appendix I

The Imbedding Theorem for Closed Manifolds

We shall give a simple proof that any closed C^r-m-manifold M $(1 \leqq r \leqq \infty)$ may be C^r-imbedded in a Euclidean space E_μ. We shall then show that if M is either open or closed, $\mu > 2m + 1$, and $2 \leqq r \leqq \omega$, then M may be projected in some direction into a subspace $E_{\mu-1}$, and hence into E_{2m+1}. The latter fact is useful also in other connections. The two facts together prove the imbedding theorem for closed manifolds with $2 \leqq r \leqq \infty$.[12]

To prove the first statement, let $U_1, U_2, \cdots$ be coordinate systems of class C^r defining M, using the closed m-cube Q: $-1 \leqq x_i \leqq 1$ $(i = 1, \cdots, m)$. Say $U_i = \theta_i(Q)$. Let Q' be the cube $-\frac{1}{2} \leqq x_i \leqq \frac{1}{2}$ and set $U_i' = \theta_i(Q')$. As M is closed, there is an s such that $U_1', \cdots, U_s'$ cover M. Let $h(x)$ be a function of class C^r of the real variable x such that

$$h(x) = \begin{cases} 0 \text{ if } x \leqq -1 \text{ or } x \geqq 1, \\ 1 \text{ if } -\frac{1}{2} \leqq x \leqq \frac{1}{2}, \end{cases}$$

$$0 < h(x) < 1 \qquad \text{for all other } x.^{13}$$

Set

$$H(x_1, \cdots, x_m) = h(x_1) \cdots h(x_m),$$

$$H_i(x_1, \cdots, x_m) = x_i H(x_1, \cdots, x_m) \qquad (i = 1, \cdots, m).$$

Then $H \equiv 1$, $H_i \equiv x_i$ in Q'. For each $i = 1, \cdots, m$ and $j = 1, \cdots, s$, set

---

[11] The use of the Dirichlet problem in this connection was suggested to me by L. Ahlfors. The n-dimensional case, with $r = \omega$, might be more easily handled as follows. Suppose N is a compact analytic manifold in E_μ, and $f(p)$ is analytic in N; we wish to extend f throughout E_μ so as to be analytic (i.e. A and B above are replaced by N and $E_\mu - N$). Perhaps one can show that for K sufficiently large, some integral equation such as

$$f(x) = T'K^n \int_N \phi(y) e^{-K^2 r_{xy}^2} dy$$

(compare AE §15) may be solved for $\phi(y)$; $f(x)$ is then analytic in E_μ.

[12] This is the proof originally discovered by the author. The method of proof can also be extended to the case of open manifolds.

[13] We may use AE, Lemma 2, or use

$$g(y) = \frac{e^{1/y}}{e^{1/y} + e^{1/(1-y)}}, \qquad h(x) = g(2x - 1) \qquad (1/2 < x < 1).$$

$$f_{0j}(p) = \begin{cases} H(\theta_j^{-1}(p)) \text{ in } U_j, \\ 0 \text{ in } M - U_j, \end{cases} \qquad f_{ij}(p) = \begin{cases} H_i(\theta_j^{-1}(p)) \text{ in } U_j, \\ 0 \text{ in } M - U_j. \end{cases}$$

This is a set of $\mu = (m + 1)s$ functions of class C^r in M. Arrange them in a definite order.

Let ϕ map each point of M into that point of E_μ with these functions as coordinates; ϕ is of class C^r. To show that ϕ is regular, take any p in M; say p is in U_j'. Project $\phi(U_j')$ onto the m-plane P of the functions $f_{1j}, \cdots, f_{mj}$, in E_μ, and let ψ be the map of U_j' into P thus defined. As $H_i \equiv x_i$ in Q', ψ maps each point $(x_1, \cdots, x_m)$ of U_j' into the point of P with the coordinates $f_{ij} = x_i$. Hence ψ is regular, and consequently ϕ is regular at p. To show that ϕ is (1-1), let p and q be distinct points of M. If they are both in some U_j', the above proof shows that $\phi(p) \neq \phi(q)$. If not, say p is in U_j' while q is not. Then $f_{0j}(p) = 1$ and $f_{0j}(q) < 1$, and again $\phi(p) \neq \phi(q)$.

In the proof of Theorem 5 we shall use the following extension of DM Lemma 13.

LEMMA 5. *If $A = A_1 + A_2 + \cdots$ is a subset of the C^1-k-manifold S, and each A_i is of zero k-extent, then $S - A$ is dense in S.*

Each A_i is nowhere dense in S (apply DM Lemma 13 to $\bar{A}_i$); the statement is a consequence of the Baire Theorem.[14]

We shall now prove the second statement.

THEOREM 5. *If M is a C^r-m-manifold $(2 \leqq r \leqq \omega)$ in E_μ, $\mu > 2m + 1$, we may imbed it by projection in a subspace $E_{\mu-1}$. If $\mu = 2m + 1$, we may project it into a local manifold in E_{2m}. The direction along which we project may in either case be taken arbitrarily near a given direction.*[15]

Let $S_{\mu-1}$ be the unit $(\mu - 1)$-sphere in E_μ. For p and q in M, $p \neq q$, there is a point $\phi(p, q)$ of $S_{\mu-1}$ whose direction from the origin is the same as that of q from p. For p in M, there is an $(m - 1)$-sphere $S(p)$ in $S_{\mu-1}$, the intersection of $S_{\mu-1}$ with the m-plane through the origin and parallel to the tangent plane to M at p. We shall find a point x arbitrarily near any x_0 in $S_{\mu-1}$ which is on no $\phi(p, q)$ or $S(p)$; this gives a direction along which we may project. Let U_1, $U_2, \cdots$ be coordinate neighborhoods in M such that each p is in arbitrarily small U_i. For each i, j such that $\bar{U}_i \cdot \bar{U}_j = 0$, let A_{ij} be the set of all points $\phi(p, q)$ for p in U_i and q in U_j; then $\sum A_{ij}$ covers all the $\phi(p, q)$. Each A_{ij} is of finite $2m$-extent and hence of zero $(\mu - 1)$-extent (DM, Lemmas 15 and 14). Let B_i be all points on all $S(p)$ for p in U_i. As M is of class C^2, B_i is a C^1-map of the product of the m-cell Q and an $(m - 1)$-sphere, and hence is of zero

---

[14] We may take open spheres G_i in S whose diameters approach 0, such that $\bar{G}_i$ is in G_{i-1} and has no points in A_i. Then there is a point p in each G_i; p is not in A.

[15] The theorem is not true for $r = 1$. For let $\phi(t)$ map the unit interval $0 \leqq t \leqq 1$ continuously into the whole of the unit sphere $S_{\mu-1}$ in E_μ. There is a curve of class C^1 in E_μ, $x = \psi(t)$, whose direction for each t is given by $\phi(t)$, with t = arc length approximately; it may be made (1-1). Joining the ends gives a C^1-1-manifold for which there is no projection as in the theorem.

$2m$-extent. There is a point x near x_0 and not on any A_{ij} or B_i, by Lemma 5. To prove the second statement in the theorem, we consider merely the B_i.

Appendix II

It seems worth while to give more details in the *proof of* DM *Lemma* 16, in particular, of the statements at the top of p. 662.

(a) If no ξ existed, there would be for each integer n an $\epsilon_n > 0$ and points x_n and y_n of A within ϵ_n of each other, whose transforms under T_{α^0} are distant apart at least $n\epsilon_n$. As A is closed and bounded, $T_{\alpha^0}(A)$ is bounded; hence $\epsilon_n \to 0$. We may therefore suppose $x_n \to x'$, $y_n \to x'$, x' in A. But this clearly contradicts the differentiability of $T_{\alpha^0}(x)$ at $x = x'$. The same proof (using α^n in place of α^0) shows that the statement holds for all α within η of α^0.

(b) A directional derivative $\partial x'/\partial \bar{\alpha}_i$ is a partial derivative after a rotation of axes: $\alpha_j = \sum a_{ji}\bar{\alpha}_i$, $\| a_{ij} \|$ an orthogonal matrix. Hence

$$\frac{\partial x'}{\partial \bar{\alpha}_i} = \sum c_j \frac{\partial x'}{\partial \alpha_j}, \qquad \sum c_j^2 = \sum_j a_{ji}^2 = 1.$$

As the $\partial x'/\partial \alpha_j$ are independent and the sets of c_j run over the unit $(h - 1)$-sphere, which is self-compact, the possible $\| \partial x'/\partial \bar{\alpha}_i \|$ have a lower bound $\beta > 0$. As in (a), we may choose β so that this holds for all α within η of α^0.

(c) To make one of $T_\alpha(A^*)$, $T_{\alpha'}(A^*)$ fail to intersect B^*, it is clearly sufficient to take α and α' such that for each x in A,

$$\Delta = \| T_{\alpha'}(x) - T_\alpha(x) \| \geqq \xi\epsilon + \epsilon.$$

If α' approaches α along a straight line lying within η of α^0, then $\lim \Delta / \| \alpha' - \alpha \|$ is a directional derivative, and hence is $\geqq \beta$. Hence, for α and α' within some ζ of $\alpha^0(\zeta \leqq \eta)$, $\Delta \geqq \frac{1}{2} \| \alpha' - \alpha \| \beta$. Now if $\| \alpha' - \alpha \| \geqq 4\xi\epsilon/\beta$, then $\Delta \geqq 2\xi\epsilon \geqq \xi\epsilon + \epsilon$.

Harvard University.

ANNALS OF MATHEMATICS
Vol. 38, No. 4, October, 1937

ON THE FIXED POINT FORMULA

BY S. LEFSCHETZ

(Received June 10, 1937)

The object of the present note is to give a new and final proof of the fixed point formula for LC spaces (= absolute neighborhood retracts). For the sake of simplicity we consider only continuous single valued transformations (= c.s.v.t.). A comparison with our earlier proof[1] will show that practically the whole burden may now be thrown on the abstract complex, the extension to abstract spaces being reduced to the barest minimum.

1. Our starting point is the fixed element formula for abstract complexes.[2] Let K be a finite complex and let E^i_p be its elements. We shall only consider rational chains and their transformations. A transformation T of the type envisaged has the form

$$TE^i_p = x^i_{p,j}E^j_p; \qquad x^i_{p,j} \text{ rational.} \tag{1.1}$$

In order that T possess some fixed element, i.e. that some E^i_p be a member of the chain TE^i_p, it is necessary and sufficient that some number $x^i_{p,i} \neq 0$ (i unsummed). A sufficient condition is then that

$$\theta = \sum (-1)^p \operatorname{trace} x_p \neq 0, \qquad x_p = \| x^i_{p,j} \| . \tag{1.2}$$

The transformations that really matter are those which permute with F, the boundary operator: $TF = FT$. This property means that T transforms a chain and its boundary into a chain and its boundary, and in particular a cycle into a cycle. It may be shown then that if $\{\gamma^i_p\}$ is a base for the p-cycles and

$$T\gamma^i_p = y^i_{p,j}\gamma^j_p, \qquad \| y^i_{p,j} \| = y_p,$$

we have trace x_p = trace y_p and hence

$$\theta = \sum (-1)^p \operatorname{trace} y_p . \tag{1.3}$$

The precise result which we need may now be explicitly formulated as

[1] See our Colloqium Lectures, *Topology*, New York (1930), Chapter VII, pp. 347, 359.

[2] Details and references regarding the author's proof for manifolds, Hopf's extension to geometric complexes and the author's proof of the same will be found in *Topology*, Chapter VI. The proof for abstract complexes is due to A. W. Tucker: Annals of Math., vol. 34 (1933) p. 238. He also introduced the transformations permutable with F. See in regard to the abstract complex: S. Lefschetz, Bull. American Math. Soc., vol. 43 (1937).

THEOREM 1. *If a chain-transformation T permutable with F has no fixed elements its invariant θ computed from* (1.3) *is zero. Therefore a sufficient condition for the presence of fixed elements is $\theta \neq 0$.*

It is this theorem which we propose to extend to LC spaces.

2. We recall briefly certain needed definitions, notably of the singular elements and of LC spaces.[3] We are dealing throughout with a basic compact metric space R. The *singular* elements of R may be introduced in two distinct manners. Under one definition the singular cells are made the basic elements and the chains, cycles, complexes defined in terms of the cells.[4] Under the other definition, which we shall adopt here, each element is considered as the c.s.v.t. of a simplicial polyhedral antecedent determined up to a simplicial homeomorphism.[5] It may be shown that any element of the first type has a subdivision of the second type so that as regards homology groups the difference between the two types is not very great.

A *semi-singular* complex K is a singular image of a closed subcomplex L of a given simplicial complex K^* which includes all the vertices of K^*. The *mesh* of a singular complex is the maximum diameter of its cells. The mesh of a semi-singular complex K associated as above with K^* is the maximum of the diameters of the realized cells and of the sum of the realized faces of the non-realized cells.

The space R is said to be LC (= locally connected) if for each $\epsilon > 0$ there exists an $\eta(\epsilon) > 0$ such that any semi-singular complex K of mesh $<\eta$ can be completed to a full singular complex of mesh $<\epsilon$.

Suppose R is an LC space and $\epsilon > 0$ is given. Choose $\delta = \eta(\epsilon)/4$, $\mu = \eta(\delta)$ and $\nu = \min(\mu/3, \delta)$. Then choose a finite set of points $\{A^i\}$ in R so that every point is within a distance ν of this set. Let Φ' be the complex having the A's for its vertices, it being agreed that any set of vertices form a simplex if their diameter is $<\mu$. We may regard Φ' as a semi-singular complex on R, in this case the closed subcomplex is the set of vertices. As its mesh is $<\mu$, it may be completed to a singular complex Φ of mesh $<\delta$.

Let τ be a point transformation defined throughout R and sending each point into one of the points A^i nearest to it. Let K be a singular complex on R of mesh $<\nu$. Consider the transformation τ applied only to the vertices of K. For any simplex of K the diameter of the set into which its vertices are mapped is $<3\nu \leqq \mu$. Hence this is a simplicial map of K into Φ'. Construct the abstract complex $\bar{K}$ composed of K and Φ' together with the prismatic cells joining each simplex of K to its image in Φ'. $\bar{K}$ may be simplicially subdivided

[3] See S. Lefschetz, Annals of Math., vol. 35 (1934), pp. 118–129 (errata in our paper: Duke Journal, vol. 1 (1935), p. 1.)

[4] See S. Lefschetz, Bull. American Math. Soc., vol. 39 (1933) pp. 124–129.

[5] See notably Alexandroff-Hopf: *Topologie*, p. 332. They call this second type of singular element "continuous".

without touching K or Φ' and without introducing further vertices.[6] In R we have the partial map of $\bar{K}$ composed of K and Φ. The mesh of this semi-singular complex is $< 3\nu + \delta \leqq \eta(\epsilon)$; hence the missing cells of $\bar{K}$ may be inserted so as to have a diameter $< \epsilon$. Since every singular complex has a subdivision of mesh $< \nu$, we see that *every singular complex has a subdivision which is ϵ-deformable into a chain of Φ.*

We shall base the homology theory of R on singular chains and cycles.[7] An immediate consequence of the above result is that every cycle of R is homologous to a cycle of Φ. Therefore the *homology groups of R are isomorphic with subgroups of the homology groups of a finite simplicial complex.* This implies that all the homology groups for dimensions above a certain n vanish.

It follows that we may find a base for the p-cycles of Φ consisting of two sets: $\{\Gamma_p^i\}$, $i = 1, 2, \cdots, r$, a maximal independent set for R and $\{\Delta_p^i\}$, $i = 1, 2, \cdots, s$ which bound in R. In terms of Betti-numbers: $r = R_p(R)$, $r + s = R_p(\Phi')$. These sets need only to be considered for $p \leqq n$. We have then explicitly certain chains C_{p+1}^i such that $F(C_{p+1}^i) = \Delta_p^i$.

3. We are now ready to take up the proof of the fixed point formula for the LC space R. Let T be a c.s.v.t. of R into itself. Its effect upon the cycles of R is described by homologies

$$T\Gamma_p^i \sim y_{p,j}^i \Gamma_p^j, \qquad y_p = \| y_{p,j}^i \|, \tag{3.1}$$

and we may therefore introduce the number

$$\theta = \sum (-1)^p \operatorname{trace} y_p . \tag{3.2}$$

θ is clearly independent of the choice of bases and hence is a function of the "homology-class" of T. We have then the following extension of Theorem 1:

THEOREM 2. *If T is a c.s.v.t. of the LC space R into itself without fixed points the invariant $\theta(T) = 0$. Therefore a sufficient condition for the presence of fixed points is $\theta \neq 0$.*

4. Suppose T is without fixed points. We may then choose ϵ so small that every point is displaced a distance $> 4\epsilon$ by T, and ν so small that any set of diameter $< \nu$ is at a distance $> \epsilon$ from its transform. This is to be in addition to the other inequalities imposed in No. 2.

Let now Ψ be the finite singular complex consisting of Φ and of the cells of

---

[6] Let $P^1, \cdots, P^r$ be the vertices of K, let $Q^i = \tau P^i$, and let $\zeta = P^h, \cdots, P^k$ be any simplex of K. The sets of vertices $P^k, \cdots, P^i Q^i \cdots Q^k$ are vertices of certain simplexes whose dimension is $\dim \zeta + 1$. Their realized faces are those in K or Φ. The LC condition enables us to insert them for all cells of K at once and the realized cells corresponding to ζ have for sum the deformation cell $D\zeta$ of ζ in the sense of *Topology* p. 78. To have $D\zeta$ oriented we must take the alternate sum of the cells.

[7] For LC spaces these cycles and their bounding relations yield the same homology groups as Vietoris cycles. See *Topology*, p. 333.

the chains C^i_{p+1} $(p = 0, 1, \cdots, n)$. We may ϵ-deform a λ-subdivision of $T\Psi$ into Φ (No. 2). Let $\overline{T}$ be the induced chain-transformation on Φ. That is to say, any cell ζ^i_p of Φ becomes first $T\zeta^i_p$ then, after subdivision and ϵ-deformation, goes into a certain new chain $\overline{T}\zeta^i_p$ of Φ. We have then

$$\overline{T}\zeta^i_p = x^i_{p,j}\zeta^j_p. \tag{4.1}$$

Since both T and the ϵ-deformation are permutable with F the same holds for $\overline{T}$. Therefore $\overline{T}\Gamma^i_p$ is a cycle of Φ and

$$F(\overline{T}C^i_{p+1}) = \overline{T}\Delta^i_p. \tag{4.2}$$

Since a deformation transforms a cycle into a homologous cycle we have

$$\overline{T}\Gamma^i_p \sim T\Gamma^i_p \text{ on } R.$$

It follows that the cycle of Φ

$$\overline{T}\Gamma^i_p - y^i_{p,j}\Gamma^j_p \sim 0 \text{ on } R$$

and hence it is $\sim$ to a linear combination of Δ's on Φ, or

$$\overline{T}\Gamma^i_p \sim y^i_{p,j}\Gamma^j_p + z^i_{p,j}\Delta^j_p \text{ on } \Phi. \tag{4.3}$$

From (4.2) follows also

$$\overline{T}\Delta^i_p \sim 0 \text{ on } \Phi. \tag{4.4}$$

Hence

$$\theta(\overline{T}) = \sum (-1)^p \text{ trace } y_p = \theta(T).$$

Now in view of the choice of $\epsilon, \lambda, \mu, \nu$, every chain at the right in (4.1) is at a distance $> \epsilon > 0$ from ζ^i_p. Therefore $\overline{T}$ has no fixed element, so that by Theorem 1, $\theta(\overline{T}) = 0 = \theta(T)$. This proves Theorem 2.

PRINCETON UNIVERSITY

Annals of Mathematics
Vol. 38, No. 4, October, 1937

ON BICOMPACT SPACES

By Eduard Čech

(Received February 3, 1937)

The theory of bicompact spaces was extensively studied by P. Alexandroff and P. Urysohn in their paper *Mémoire sur les espaces topologiques compacts,* Verhandlingen der Kon. Akademia Amsterdam, Deel XIV, No. 1, 1929; I shall refer to this paper with the letters AU. An important result was added by A. Tychonoff in his paper *Über die topologische Erweiterung von Räumen,* Math. Annalen 102, 1930, who proved that complete regularity is the necessary and sufficient condition for a topological space to be a subset of some bicompact Hausdorff space. As a matter of fact, Tychonoff proves more, viz. that, given a completely regular space S, there exists a bicompact Hausdorff space $\beta(S)$ such that (i) S is dense in $\beta(S)$, (ii) any bounded continuous real function defined in the domain S admits of a continuous extension to the domain $\beta(S)$. It is easily seen that $\beta(S)$ is uniquely defined by the two properties (i) and (ii). The aim of the present paper is chiefly the study of $\beta(S)$.

The paper is divided into four chapters. In chapter I, I briefly resume some well known definitions adding a few simple remarks. In particular I show that an *arbitrary* topological space S determines a completely regular space $\rho(S)$ such that a good deal of topology of S reduces to the topology of $\rho(S)$, this being true in particular for the theory of real valued continuous and Baire functions. Chapter II contains the theory of the bicompact space $\beta(S)$ mentioned above. Here I shall recall only a few results of chapter II. First, if the space S is *normal,* then $\beta(S)$ may be defined without any reference to continuous real function since property (ii) may be replaced by the following: if two closed subsets of S have no common point, then their closures in $\beta(S)$ have no common point either. Second, if the space S satisfies the first countability axiom, then S is completely determined by $\beta(S)$, S being simply the set of all points of $\beta(S)$ where the first countability axiom holds true. This implies that in this case (embracing the case of metrizable spaces) the whole topology of S may be reduced to the topology of the bicompact space $\beta(S)$. Hence it is evident that it is highly desirable to carry further the study of bicompact spaces and in particular of $\beta(S)$. Of course it must be emphasized that $\beta(S)$ may be defined only formally (not constructively) since it exists only in virtue of Zermelo's theorem. If I denotes the space of integer numbers, then I think it is impossible to determine effectively (in the sense of Sierpiński) a point of $\beta(I) - I$. I was even unable to determine the cardinal number of $\beta(I)$. (The paper contains several other unsolved problems.) The space $\beta(I) - I$ furnishes incidentally a positive solution of a problem proposed by Alexandroff and Urysohn (AU, p. 54:

Existe-t-il un espace bicompact ne contenant aucun point (κ)? The authors write in this connection: La résolution affirmative de ce problème nous donnerait un exemple des espace bicompacts d'une nature toute différente de celle des espaces connus jusqu'à présent). In chapter III, I call a completely regular space S *topologically complete* if S is a G_δ in $\beta(S)$. The reason for this designation lies in the fact that, if S is metrizable, it has this property if and only if it is homeomorphic with a metric complete space. The proof is an easy adaptation of Hausdorff's well known proof of the theorem that a G_δ in a metric complete space is a homeomorph of a metric complete space. In chapter IV, I consider *locally normal* spaces and I prove that a locally normal space S is always an open subset of some normal space. This was of course to be expected but I think it would be difficult to prove without the theory of $\beta(S)$.

I

A set S is called a *topological space* (and its elements are called *points*) if there is given a class $\mathfrak{F}$ of subsets of S (called *closed* subsets of S) such that (1) the whole space S and the vacuous set 0 are closed, (2) the intersection of any family of closed sets is closed, (3) the sum of two closed sets is closed. A set $G \subset S$ is called *open*, if the complementary set $S - G$ is closed. A *neighborhood* of a set $A \subset S$ (A may consist of a single point) is an open set containing A.

The intersection of all closed sets containing a given set A is called the *closure* of A and is denoted by $\bar{A}$. The closure operation has the following properties: (1) $\bar{0} = 0$, (2) $A \subset \bar{A}$, (3) $\overline{A + B} = \bar{A} + \bar{B}$, (4) $\bar{\bar{A}} = \bar{A}$. Conversely, it is possible to define the general notion of a topological space starting with an operation $\bar{A}$ subject only to conditions (1)-(4) and defining closed sets by the condition $\bar{A} = A$.

An *open base* of a topological space S is a class $\mathfrak{B}$ of open sets such that any open set is the sum of some of the elements of $\mathfrak{B}$. The class $\mathfrak{H}$ of *all* open sets is a particular open base. Any open base $\mathfrak{B}$ has the following properties: (1) given a point $x \in S$, there exists a $U \in \mathfrak{B}$ such that $x \in U$, (2) given a point $x \in S$ and two sets U and V such that $U \in \mathfrak{B}$, $V \in \mathfrak{B}$, $x \in UV$, there exists a set W such that $W \in \mathfrak{B}$, $x \in W$, $W \subset UV$. Conversely it is possible (and the possibility is utilized very frequently in practice) to define a topological space starting with a class $\mathfrak{B}$ subject only to condition (1) and (2); the closure $\bar{A}$ of a set $A \subset S$ consists then of all the points x such that

$$U \in \mathfrak{B},\ x \in U \text{ implies } UA \neq 0.$$

A fixed subset T of a topological space S is always considered as a topological space, defining a set $A \subset T$ to be *relatively closed* (i.e. closed in the space T) whenever A is the intersection of T with some closed subset of S. A set $A \subset T$ is *relatively open* whenever A is the intersection of T with some open subset of S. The *relative closure* of a set $A \subset T$ is the intersection $T\bar{A}$ of T with the closure of A in the space S. Any open base $\mathfrak{B}$ of S determines an open base $\mathfrak{B}_0$ of T; the elements of $\mathfrak{B}_0$ are the intersections of T with the elements of $\mathfrak{B}$.

A *mapping* f of a topological space S_1 into a topological space S_2 is an operation attaching to each point $x \in S_1$ a definite point $f(x) \in S_2$; we always suppose that, given any point $y \in S_2$, there exists at least one point $x \in S_1$ such that $f(x) = y$. The space S_1 is the *domain* of f, S_2 is its *range*. The *image* $f(A)$ of a set $A \subset S_1$ is the set of all points $f(x)$, x running over A. The *inverse image* $f^{-1}(B)$ of a set $B \subset S_2$ is the set of all points $x \in S_1$ such that $f(x) \in B$. The mapping f is *one-to-one* if

$$x_1 \in S_1,\ x_2 \in S_1,\ x_1 \neq x_2 \text{ implies } f(x_1) \neq f(x_2).$$

If f is one-to-one, then the inverse operation f^{-1} is a one-to-one mapping of S_2 into S_1. The mapping f will be called a *function* if its range consists of real numbers. The function f is *bounded* if its range is a bounded set.

The mapping f is called *continuous at a point* $x \in S_1$ if, given any neighborhood V of $f(x)$, there exists a neighborhood U of x such that $f(U) \subset V$. f is called *continuous* (simply) if it is continuous at any point $x \in S_1$. f is called *homeomorphic* if it is one-to-one and if both f and f^{-1} are continuous. f is continuous, if and only if the inverse image of any closed subset of S_2 is a closed subset of S_1.

A set $A \subset S$ is callea a G_δ-set if there exists a countable sequence $\{G_n\}$ of open sets such that $A = \prod_1^\infty G_n$; A is called an F_σ-set if there exists a countable sequence $\{F_n\}$ of closed sets such that $A = \sum_1^\infty F_n$. The complement of a G_δ-set is an F_σ-set and vice-versa.

S is called a *Kolmogoroff space*[1] if the closures of any two distinct points are distinct. (Kol T_0) S is called a *Riesz space*[2] if any single point is closed. S is a Riesz space if and only if the intersection of all the neighborhoods of any point x consists of x only. (Riesz T_1) S is called a *Hausdorff space* if the intersection of the closures of all the neighbhorhoods of any point x consists of x only. (Haus T_2) Any Riesz space is a Kolmogoroff space. Any Hausdorff space is a Riesz space. Any subset of a Kolmogoroff space is a Kolmogoroff space. Any subset of a Riesz space is a Riesz space. Any subset of a Hausdorff space is a Hausdorff space. Let $\mathfrak{B}$ be any open base of S. S is a Kolmogoroff space if and only if, given two distinct points x and y, there exists a set $U \in \mathfrak{B}$ containing precisely one of the points x and y. S is a Riesz space if and only if, given two distinct points x and y, there exists a set $U \in \mathfrak{B}$ containing x and not containing y. S is a Hausdorff space if and only if, given two distinct points x and y, there exist sets U and V such that $U \in \mathfrak{B}$, $V \in \mathfrak{B}$, $x \in U$, $y \in V$, $UV = 0$.

Now we proceed to prove that *the theory of general topological spaces* (in the sense precised above) *can be completely reduced to the theory of Kolmogoroff spaces.* Let S be a topological space. Two points $x \in S$ and $y \in S$ will be called equivalent (for the time being) if $\bar{x} = \bar{y}$. Let F be any closed subset of S and let x and y be two equivalent points; if $x \in F$, then $\bar{x} \subset F$, since F is closed, but $y \in \bar{y}$ and $\bar{y} = \bar{x}$, so that $y \in F$. It follows that any closed subset of S consists of complete

[1] See P. Alexandroff and H. Hopf, *Topologie* I, p. 58.

[2] See G. Birkhoff, *On the combination of topologies*, Fund. Math. 26, p. 162.

classes of mutually equivalent points. Now let us attach to each point $x \in S$ a new symbol $\tau(x)$ chosen in such manner that $\tau(x) = \tau(y)$ if and only if x and y are equivalent; let us call S_0 the set of the symbols $\tau(x)$, so that τ is a mapping of S into S_0. A set $A_0 \subset S_0$ will be considered as closed if and only if its inverse image $\tau^{-1}(A_0)$ is a closed subset of S. It is evident that S_0 is a topological space and that τ is a continuous mapping. Further it is evident that for any set $A \subset S$ we have $\tau(\bar{A}) \doteq \overline{\tau(A)}$; in particular $\tau(\bar{x}) = \overline{\tau(x)}$ for any $x \in S$. If $\tau(x) \neq \tau(y)$, we have $\bar{x} \neq \bar{y}$; since the sets $\bar{x}$ and $\bar{y}$ are closed, it easily follows that $\tau(\bar{x}) \neq \tau(\bar{y})$, or $\overline{\tau(x)} \neq \overline{\tau(y)}$, so that S_0 is a Kolmogoroff space. Conversely, let S_0 be a Kolmogoroff space. Let τ be a mapping of a set S into S_0. Let us call closed in S the inverse image of any closed subset of S_0. Then S is the most general topological space and τ has the previous meaning. Evidently the topology of S is quite completely described by that of S_0.

S is called a *regular space* if it is a Kolmogoroff space having the following property: given a neighborhood U of a point x, there exists a neighborhood V of x such that $\bar{V} \subset U$.[3] We shall prove that any regular space S is a Hausdorff space.[4] Let x and y be two distinct points of S. If we had both $x \in \bar{y}$ and $y \in \bar{x}$, it would follow, since $\bar{x}$ and $\bar{y}$ are closed, that $\bar{x} \subset \bar{y}$ and $\bar{y} \subset \bar{x}$, i.e. $\bar{x} = \bar{y}$, which is impossible. The argument being symmetrical, we may suppose that x does not belong to $\bar{y}$, so that $S - \bar{y}$ is a neighborhood of x. Hence there exists a neighborhood U of x such that $\bar{U} \subset S - \bar{y}$. Putting $V = S - \bar{U}$, we have two open sets U and V such that $x \in U$, $y \in V$, $UV = 0$, so that S is a Hausdorff space.

Any subset of a regular space is a regular space.

S is called a *completely regular space* if it is a Kolmogoroff space having the following property: given a closed set F and a point $a \in S - F$, there exists a continuous function f (in the domain S) such that $f(a) = 0$ and $f(x) = 1$ for any $x \in F$.[5] It is easy to see that a completely regular space is regular and that any subset of a completely regular space is a completely regular space.

Now we shall start with an arbitrary topological space S and we shall attach to it a uniquely defined completely regular space $\rho(S)$ in such manner that a great deal of topology of S may be reduced to that of $\rho(S)$. Two points x and y of S will be called equivalent (for the time being) if $f(x) = f(y)$ for every continuous function f (in the domain S). To each point $x \in S$ let us attach a new symbol $\rho(x)$ chosen in such a manner that $\rho(x) = \rho(y)$ if and only if x and y are equivalent;[6] let us call S_1 the set of all the symbols $\rho(x)$, so that ρ is a mapping of S into $S_1 = \rho(S)$. We shall introduce a topology in S_1 by defining an open

---

[3] The neighborhoods may here be restricted to a given open base of S.

[4] This is usually done assuming *a priori* that S is a Riesz space; for this point I am indebted to Dr. K. Koutský.

[5] We may assume that $0 \leqq f(x) \leqq 1$ for every $x \in S$, since we could replace f with φ by defining $\varphi(x) = f(x)$ if $0 \leqq f(x) \leqq 1$, $\varphi(x) = 0$ if $f(x) < 0$, and $\varphi(x) = 1$ if $f(x) > 1$.

[6] It is evident that $\tau(x) = \tau(y)$ implies $\rho(x) = \rho(y)$, but of course we may restrict ourselves to Kolmogoroff spaces.

base $\mathfrak{B}$ for S_1. An element $[f, I]$ of $\mathfrak{B}$ will be defined by a continuous function f in the domain S and an open interval I, $[f, I]$ consisting of the points $\rho(x)$ of S_1 such that $f(x) \in I$. To prove that S_1 is a topological space we have to verify two things. First, for any $a \in S$, there evidently exists an $[f, I]$ containing $\rho(a)$. Second, let $\rho(a)$ belong both to $[f_1, I_1]$ and to $[f_2, I_2]$; we have to prove that there exists an $[f, I]$ such that $\rho(a) \in [f, I]$ and $[f, I] \subset [f_1, I_1] \cdot [f_2, I_2]$. There exists a number $\varepsilon > 0$ such that, for $i = 1$ and for $i = 2$, the interval $f_i(a) - \varepsilon < t < f_i(a) + \varepsilon$ is a subset of I_i. It is easy to see that we may put $f(x) = | f_1(x) - f_1(a) | + | f_2(x) - f_2(a) |$, choosing I to be the interval $-\varepsilon < t < \varepsilon$. Hence S_1 is a topological space.

Since the topology of S_1 was defined by means of *continuous* functions in the domain S, it is easy to see that ρ is a continuous mapping of S into S_1 so that, if φ is any continuous function in the domain S_1, $f(x) = \varphi[\rho(x)]$ is a continuous function in the domain S. Moreover, in our case the converse is also true: *any continuous function in the domain S has the form $f(x) = \varphi[\rho(x)]$, φ being a continuous function in the domain S_1.*

If $\rho(a)$ and $\rho(b)$ are two distinct points of S_1, then there exists a continuous function f in the domain S such that $f(a) \neq f(b)$. There exist two disjoined open intervals I_1 and I_2 such that $f(a) \in I_1$ and $f(b) \in I_2$. Then $[f, I_1]$ and $[f, I_2]$ are two disjoined open subsets of S_1 and $\rho(a) \in [f, I_1]$, $\rho(b) \in [f, I_2]$. It follows that S_1 is a Hausdorff space. As a matter of fact, S_1 is a completely regular space. Let Φ be a closed subset of S_1 not containing the point $\rho(a)$. There exists an $[f, I]$ such that $\rho(a) \in [f, I] \subset S_1 - \Phi$; we may suppose that I consists of all numbers t such that $| t - f(a) | < \varepsilon (\varepsilon > 0)$. If $| f(x) - f(a) | \geqq \varepsilon$, put $g(x) = 1$; if $| f(x) - f(a) | < \varepsilon$, put $g(x) = \varepsilon^{-1} . | f(x) - f(a) |$. Then g is a continuous function in the domain S, so that there exists a continuous function φ in the domain S_1 such that $g(x) = \varphi[\rho(x)]$. It is easy to see that $\varphi[\rho(a)] = 0$ and $\varphi(x) = 1$ for each $x \in \Phi$.

Let F be a closed subset of S. We shall prove that *a necessary and sufficient condition for the set $\rho(F)$ to be closed in S_1 is that for any point*

$$a \in S - \rho^{-1}[\rho(F)]$$

there exists a continuous function f in the domain S such that $f(a) = 0$ and $f(x) = 1$ for each $x \in F$. First suppose the condition satisfied. If $\rho(F)$ were not closed in S_1, we could choose a point a such that

$$\rho(a) \in \overline{\rho(F)} - \rho(F).$$

Since $\rho(a) \in S_1 - \rho(F)$, there would exist a continuous function f in the domain S such that $f(a) = 0$ and $f(x) = 1$ for each $x \in F$. There would exist a continuous function φ in the domain S_1 such that $f(x) = \varphi[\rho(x)]$. For $x \in \rho(F)$ we would have $\varphi(x) = 1$; since φ is continuous, it easily follows that $\varphi(x) = 1$ for $x \in \overline{\rho(F)}$, in particular $\varphi[\rho(a)] = 1$, i.e. $f(a) = 1$, which is a contradiction. Secondly, suppose $\rho(F)$ closed in S_1. Let $a \in S - \rho^{-1}[\rho(F)]$. Then $\rho(a) \in S_1 - \rho(F)$. Since S_1 is completely regular, there exists a continuous function φ in the domain

S_1 such that $\varphi[\rho(a)] = 0$ and $\varphi(x) = 1$ for each $x \in \rho(F)$. Putting $f(x) = \varphi[\rho(x)]$, we have a continuous function f in the domain S such that $f(a) = 0$ and $f(x) = 1$ for each $x \in F$.

As a corollary, we obtain that, if the space S itself is completely regular, the mapping ρ is homeomorphic.

The following property is characteristic for completely regular spaces S: *Let σ be a continuous mapping of S into a topological space R such that each continuous function f in the domain S has the form $f(x) = \varphi[\sigma(x)]$, φ being a continuous function in the domain R. Then the mapping σ is homeomorphic.* The property cannot be true if S is not completely regular, as is seen by putting $\sigma = \rho$. Hence suppose that S is completely regular. If $a \in S$, $b \in S$, $a \neq b$, there exists a continuous function f in the domain S such that $f(a) \neq f(b)$; since $f(x) = \varphi[\sigma(x)]$, we have $\sigma(a) \neq \sigma(b)$, i.e. the mapping σ is one-to-one. It remains to show that if F is a closed subset of S the set $\sigma(F)$ is closed in R. If $\sigma(F)$ is not closed, there exists a point $a \in S$ such that

$$\sigma(a) \in \overline{\sigma(F)} - \sigma(F).$$

There exists a continuous function f in the domain S such that $f(a) = 0$ and $f(x) = 1$ for each $x \in F$. We may put $f(x) = \varphi[\sigma(x)]$ and we have $\varphi[\sigma(a)] = 0$ and $\varphi(x) = 1$ for each $x \in \sigma(F)$. Since φ is continuous, we must have $\varphi(x) = 1$ for each $x \in \overline{\sigma(F)}$, hence for $x = a$, which is a contradiction.

Consider the following three properties of a topological space S: (1) If F_1 and F_2 are two closed sets such that $F_1F_2 = 0$, there exist two open sets G_1 and G_1 such that $F_1 \subset G_1$, $F_2 \subset G_2$, $G_1G_2 = 0$. (2) If F_1 and F_2 are two closed sets such that $F_1F_2 = 0$, there exists a continuous function f in the domain S such that $f(x) = 0$ for each $x \in F_1$ and $f(x) = 1$ for each $x \in F_2$.[5] (3) If F is a closed set and if φ is a bounded[7] continuous function in the domain F, there exists a continuous function f in the domain S such that $f(x) = \varphi(x)$ for each $x \in F$. It is easily seen that (2) is formally stronger than (1) and that (3) is formally stronger than (2). But Urysohn proved[8] that all three properties are equivalent to one another. A space having these properties is called *normal*. Property (2) shows that a normal Riesz space is a completely regular space (hence a regular space, therefore a Hausdorff space).

If the space S is normal, then $\rho(S)$ is normal as well. Let Φ_1 and Φ_2 be two closed subsets of $\rho(S)$ such that $\Phi_1\Phi_2 = 0$. Then $F_1 = \rho^{-1}(\Phi_1)$ and $F_2 = \rho^{-1}(\Phi_2)$ are two closed subsets of S such that $F_1F_2 = 0$. Since S is normal, there exists a continuous function f in the domain S such that $f(x) = 0$ for each $x \in F_1$ and $f(x) = 1$ for each $x \in F_2$. There exists a continuous function φ in the domain $\rho(S)$ such that $f(x) = \varphi[\rho(x)]$. Evidently $\varphi(x) = 0$ for each $x \in \Phi_1$ and $\varphi(x) = 1$ for each $x \in \Phi_2$.

If the space S is normal, then for $a \in S$, $b \in S$ we have $\rho(a) = \rho(b)$ if and only if

---

[7] It is easy to prove that the word *bounded* may be omitted.

[8] P. Urysohn, *Über die Mächtigkeit zusammenhängender Mengen*, Math. Annalen 94, 1925.

$\bar{a}\cdot\bar{b} \neq 0$. Suppose first that $c \in \bar{a}\cdot\bar{b}$. If f is a continuous function in the domain S, it is easy to see that $f(a) = f(c) = f(b)$, whence $\rho(a) = \rho(b)$. Secondly, suppose that $\bar{a}\cdot\bar{b} = 0$. Since S is normal, there exists a continuous function f in the domain S such that $f(x) = 0$ for each $x \in \bar{a}$ and $f(x) = 1$ for each $x \in \bar{b}$, whence $f(a) = 0, f(b) = 1$.

If the space S is normal and if F is a closed subset of S, then $\rho(F)$ is a closed subset of $\rho(S)$. Let $a \in S - \rho^{-1}[\rho(F)]$. For $x \in F$ we have $\rho(a) \neq \rho(x)$, whence $\bar{a}\cdot\bar{x} = 0$; therefore $\bar{a}\cdot F = 0$. Hence there exists a continuous function f in the domain S such that $f(x) = 1$ for each $x \in F$ and $f(x) = 0$ for each $x \in \bar{a}$, in particular $f(a) = 0$. We know that this implies that $\rho(F)$ is closed in $\rho(S)$.

The last two theorems show that, if S is normal, the space $\rho(S)$ and its topology may be completely described without any explicit reference to continuous functions: The space $\rho(S)$ consists of symbols $\rho(x)$ attached to single points $x \in S$, $\rho(x)$ and $\rho(y)$ being identical if and only if $\bar{x}\cdot\bar{y} \neq 0$; and a set $\Phi \subset \rho(S)$ is closed in $\rho(S)$ if and only if the set $\rho^{-1}(\Phi)$ is closed in S. It is an interesting problem to give a similar description of $\rho(S)$ in the general case.

If the space S is normal, then a necessary and sufficient condition for a set $A \subset S$ to be both closed and a G_δ is the existence of a continuous function f in the domain S such that $f(x) = 0$ if and only if $x \in A$. Suppose first that such a function f exists. Then $A = \{f(x) = 0\}$ is a closed set and $G_n = \{\,|f(x)| < 1/n\}$ are open sets and $A = \prod G_n$. Conversely let $A = \bar{A} = \prod G_n$, G_n being open. Since S is normal, there exist continuous functions f_n in the domain S such that $f_n(x) = 0$ for $x \in A$, $f_n(x) = 1$ for $x \in S - G_n$, $0 \leqq f_n(x) \leqq 1$ for $x \in S$. It is sufficient to put $f(x) = \sum 2^{-n}\cdot f_n(x)$.

A point x of a topological space S is called a *complete limit point* of a set $A \subset S$ if, for any neighborhood U of x, the cardinal number of the set AU is equal to the cardinal number of the set A. A family $\mathfrak{C}$ of subsets of S is called *monotonic* if for any two sets $A \in \mathfrak{C}$, $B \in \mathfrak{C}$ we have either $A \subset B$ or $B \subset A$. A family $\mathfrak{C}$ of subsets of S is called a *covering* of S if each point of S belongs to some set of $\mathfrak{C}$.

Consider the following three properties of a topological space S: (1) Every infinite subset possesses at least one complete limit point. (2) A monotonic family of non-vacuous closed subsets has a non-vacuous intersection. (3) Any covering of S consisting of open sets contains a finite covering of S. It is known that all three properties are equivalent to one another.[9] A space having these properties is called *bicompact*. It is known that *a bicompact Hausdorff space is normal*[10] (hence completely regular). *A closed subset of a bicompact space is a bicompact space.* Conversely, *a bicompact subset of a Hausdorff space is closed.*[11] It easily follows that *a one-to-one continuous mapping of a bicompact Hausdorff space is homeomorphic.*

Let $\{S_\iota\}$ be a family of sets; the subscript ι runs over an arbitrarily given set I. The cartesian product $\mathfrak{P}_\iota S_\iota$ of the family $\{S_\iota\}$ is the set of all families $x = \{x_\iota\}$,

[9] AU, p. 8.
[10] AU, p. 26.
[11] AU, p. 47.

each x_ι belonging to S_ι. The x_ι's are called the coordinates of x. If every S_ι is a topological space, we introduce a topology into $S = \mathfrak{P}_\iota S_\iota$ by means of the following open base $\mathfrak{B}$: The elements of $\mathfrak{B}$ are sets of the form $\mathfrak{P}_\iota G_\iota$ where (1) each G_ι is an open subset of S_ι, (2) $G_\iota = S_\iota$ except for a finite number of subscripts ι. It is easy to see that S is a Kolmogoroff space, a Riesz space, a Hausdorff space, a regular space, a completely regular space, if and only if *every* factor space S_ι belongs to the corresponding category of spaces. If S is normal, every S_ι is normal as well; but the converse is false.

The cartesian product $S = \mathfrak{P}_\iota S_\iota$ *of any family of bicompact spaces ia a bicompact space.* Using Zermelo's theorem, we may suppose that the set I consists of all ordinal numbers less than a given ordinal number. Let there be given an infinite subset A of S. We have to construct a complete limit point $z = \{z_\iota\}$ of S. According to the way the topology of S was introduced, it is sufficient to construct the coordinates z_ι by transfinite induction, choosing each $z_\iota \epsilon S_\iota$ in such a way that it have the following property π_ι : If there is given a finite number of subscripts $\iota_n \leqq \iota$ and, for each ι_n, a neighborhood G_n of z_{ι_n} (in the space S_{ι_n}), then the cardinal number of the intersection of A with the set of those points $x = \{x_\iota\}$ for which $x_{\iota_n} \epsilon G_n$ (for each of the given subscripts ι_n) is equal to the cardinal number of A. We need only prove that the definition of the z_ι's by transfinite induction may be carried through. Hence suppose that, for a definite value $\lambda \epsilon I$, the points z_ι (with property π_ι) having already been constructed for $\iota < \lambda$, it is impossible to choose $z_\lambda \epsilon S_\lambda$ with property π_λ. Then, for every point $y_\lambda \epsilon S_\lambda$, there exist: a neighborhood $T(y_\lambda)$ of the point y_λ (in the space S_λ), a finite (perhaps vacuous) set $M(y_\lambda)$ of subscripts $\iota < \lambda$ and, for each $\iota \epsilon M(y_\lambda)$, a neighborhood $G(z_\iota, y_\lambda)$ of the point z_ι (in the space S_ι) such that the cardinal number of the set $A \cdot H(y_\lambda) \cdot K(y_\lambda)$ is less than the cardinal number of A, where $H(y_\lambda)$ is the set of all points $x = \{x_\iota\}$ for which $x_\lambda \epsilon T(y_\lambda)$ and $K(y_\lambda)$ is the set of all points $x = \{x_\iota\}$ for which $x_\iota \epsilon G(z_\iota, y_\lambda)$ for every $\iota \epsilon M(y_\lambda)$. Since the space S_λ is bicompact, there exists a finite set of points $y_\lambda^{(i)} \epsilon S_\lambda (1 \leqq i \leqq m < \infty)$ such that

$$(1) \qquad \sum_{i=1}^{m} T(y_\lambda^{(i)}) = S_\lambda.$$

The cardinal number of the set

$$(2) \qquad \sum_{i=1}^{m} A \cdot H(y_\lambda^{(i)}) \cdot K(y_\lambda^{(i)})$$

is less than the cardinal number of A. On the other hand, it follows from (1) that

$$\sum_{i=1}^{m} H(y_\lambda^{(i)}) = S$$

so that the set (2) contains the set

$$(3) \qquad A \cdot \prod_{i=1}^{m} K(y_\lambda^{(i)}).$$

It follows that the cardinal number of the set (3) is less than the cardinal number of A. But it is easy to see that this is in contradiction with property π_μ, choosing $\mu < \lambda$ and $\mu \geqq \iota$ for every $\iota \in \sum_i M(y_\lambda^{(i)})$.

II

Since a bicompact Hausdorff space is completely regular, every subset of a bicompact Hausdorff space is also completely regular. Following Tychonoff, we shall prove conversely that *every completely regular space is a subset of some bicompact Hausdorff space.*

Let S be given completely regular space. Let T denote the interval $0 \leqq t \leqq 1$. Let Φ denote the set of all continuous functions f in the domain S such that $f(S) \subset T$. Choose a set I having the same potency as the set Φ, so that there exists a one-to-one mapping of I into Φ; let f_ι be the function corresponding to $\iota \in I$. For $\iota \in I$, put $T_\iota = T$ and let R be the cartesian product $\mathfrak{P}_\iota T_\iota$. Since every T_ι is a bicompact Hausdorff space, R is also a bicompact Hausdorff space. For any $x \in S$, put $g(x) = \xi = \{\xi_\iota\} \in R$, where $\xi_\iota = f_\iota(x)$. Then g is a mapping of the space S into the space $S^* = g(S) \subset R$. It is easy to see that the mapping g is homeomorphic. For $\iota \in I$ and $\xi \in R$, put $\varphi_\iota(\xi) = \xi_\iota$. Then φ_ι is a continuous function in the domain R such that $\varphi_\iota(R) = T$. Moreover, we see that $\varphi_\iota[g(x)] = f_\iota(x)$ for $x \in S$.

If S is a completely regular space, let $\beta(S)$ designate any topological space having the following four properties: (1) *$\beta(S)$ is a bicompact Hausdorff space,* (2) $S \subset \beta(S)$, (3) *S is dense in $\beta(S)$* (i.e. the closure of S in the space $\beta(S)$ is the whole space $\beta(S)$), (4) *every bounded continuous function f in the domain S may be extended*[12] *to the domain $\beta(S)$* (i.e. there exists a continuous function φ in the domain $\beta(S)$ such that $\varphi(x) = f(x)$ for every $x \in S$).

The space $\beta(S)$ exists for every completely regular S. Using the above notation, we easily see that the closure of S^* in the space R has the properties (1)-(4) relatively to S^*, so that $\beta(S^*)$ exists. Since S and S^* are homeomorphic, $\beta(S)$ exists as well.

Given a completely regular space S, the space $\beta(S)$ is essentially unique. More precisely: If B_1 and B_2 both have properties (1)-(4) of $\beta(S)$, then there exists a homeomorphic mapping h of B_1 into B_2 such that $h(x) = x$ for each $x \in S$. This is but a particular case of the following theorem: *Let S be a completely regular space. Let B be a space having properties* (1)-(3) of $\beta(S)$ (but not necessarily property (4)). *Then there exists a continuous mapping h of $\beta(S)$ into B such that*: (i) *$h(x) = x$ for each $x \in S$,* (ii) *$h[\beta(S) - S] = B - S$. The mapping h is one-to-one (and consequently homeomorphic) if and only if B also possesses property* (4). Let I, T, R, g and S^* have the above meaning. Divide the set I into two disjoined subsets I_1 and I_2, putting $\iota \in I_1$ if and only if the continuous function f_ι may be extended to the domain B. Let R_1 denote the cartesian product $\mathfrak{P}_\iota T_\iota$ where ι runs over I_1 and $T_\iota = T$ for each ι. For any $x \in B$, put $g_1(x) = \xi = \{\xi_\iota\}_{\iota \in I_1} \in R_1$, where $\xi_\iota = \varphi_\iota(x)$, φ_ι being the extension of f_ι to the domain B.

[12] It follows easily from property (3) that the extended function is uniquely defined by f.

Then g_1 is a homeomorphic mapping of the space B into the space $B^* = g_1(B) \subset R_1$, just as g was a homeomorphic mapping of S into the space S^*. For any point $\xi = \{\xi_\iota\}_{\iota \in I} \in R$, put $k(\xi) = \{\xi_\iota\}_{\iota \in I_1} \in R_1$. Evidently k is a continuous mapping of R into R_1. For $x \in S$, it is easy to see that $k[g(x)] = g_1(x)$ so that $k(S^*) \subset B^*$. Since k is continuous, it follows that $k(\overline{S^*}) \subset \overline{B^*}$, where $\overline{S^*}$ is the closure of S^* in the space R and $\overline{B^*}$ is the closure of B^* in the space R_1. Since B^* is a homeomorph of B, B^* is a bicompact Hausdorff space, whence $\overline{B^*} = B^*$. Therefore $k(\overline{S^*}) \subset B^*$, i.e. k defines a continuous mapping k_0 of $\overline{S^*}$ into a subset of B^*. Since $\overline{S^*}$ was homeomorphic with $\beta(S)$, and B^* was homeomorphic with B, k_0 defines a continuous mapping h of $\beta(S)$ into a subspace $h[\beta(S)]$ of B; evidently $h(x) = x$ for every $x \in S$. The space $h[\beta(S)]$, as a continuous image of the bicompact space $\beta(S)$, must be bicompact. It follows that $h[\beta(S)]$ is closed in B. On the other hand, $h[\beta(S)] \supset S$ must be dense in B. Therefore, $h[\beta(S)] = B$, i.e., h is a continuous mapping of $\beta(S)$ into B. If B possesses property (4) of $\beta(S)$, we have $I_1 = I$, whence $R_1 = R$ and k is the identity. This readily implies that the mapping h is homeomorphic.

Returning to the general case, we still have to prove that $h[\beta(S) - S] = B - S$. Of course $h[\beta(S) - S] \supset B - S$. It remains to arrive at a contradiction in supposing the existence of a point $b \in \beta(S) - S$ such that $a = h(b) \in S$. Since $\beta(S)$ is a bicompact Hausdorff space, it is completely regular. Hence there exists a continuous function φ in the domain $\beta(S)$ such that $\varphi(a) = 0$, $\varphi(b) = 1$. Let Q be the set of all points $x \in S$ such that $\varphi(x) \geqq \frac{1}{2}$. Then Q is a closed subset of S, so that there exists a closed subset P of the space B such that $Q = SP$. Since B is a bicompact Hausdorff space, it is completely regular. Hence there exists a continuous function ψ in the domain B such that $\psi(a) = 0$, $\psi(x) = 1$ for each $x \in P$ and $0 \leqq \psi(x) \leqq 1$ for each $x \in B$. From property (4) of $\beta(S)$ it follows that there exists a continuous function χ in the domain $\beta(S)$ such that $\chi(x) = \psi(x)$ for each $x \in S$, whence $\chi(a) = 0$. Since h is a continuous mapping of $\beta(S)$ into B, $\psi[h(x)]$ is a continuous function in the domain $\beta(S)$. The set C of all points $x \in \beta(S)$ such that $\psi[h(x)] = \chi(x)$, is closed in $\beta(S)$ and contains the set S which is dense in $\beta(S)$; therefore $C = \beta(S)$, whence $\chi(b) = \psi[h(b)] = \psi(a) = 0$. The set D of all points $x \in \beta(S)$ such that both $\varphi(x) > \frac{1}{2}$ and $\chi(x) < \frac{1}{2}$ is open in $\beta(S)$ and is not vacuous, since $b \in D$. Since S is dense in $\beta(S)$, there exists a point $c \in S \cdot D$. Since $c \in D$, we have $\chi(c) < \frac{1}{2}$; since $c \in S$, we have $\chi(c) = \psi(c)$. Therefore $\psi(c) < \frac{1}{2}$ so that $c \in S \cdot (B - P) = S - Q$. From the definition of Q it follows that $\varphi(c) < \frac{1}{2}$; since $c \in D$, this is a contradiction.

Two subsets A_1 and A_2 of a topological space S will be called *completely separated* if there exists a continuous function f in the domain S such that $f(x) = 0$ for each $x \in A_1$ and $f(x) = 1$ for each $x \in A_2$.[5] It is easy to see that A_1 and A_2 are completely separated if and only if the closed sets $\bar{A}_1$ and $\bar{A}_2$ are completely separated. We know that S is completely regular if and only if any single point x and any closed set not containing x are always completely separated. We know that S is normal if and only if two closed sets without common points are always completely separated.

Let S be a completely regular space. We characterized the space $\beta(S)$ by the properties (1)-(4) given above. We will now show that $\beta(S)$ may be also characterized by the properties (1), (2), (3) and (4′), where (4′) means the following: *If A_1 and A_2 are two completely separated subsets of S, then the closures of A_1 and A_2 in the space $\beta(S)$ are disjoint.* Suppose first that A_1 and A_2 are two completely separated subsets of S. Then there exists a continuous function f in the domain S such that $f(x) = 0$ for each $x \in A_1$ and $f(x) = 1$ for each $x \in A_2$. We may suppose that $0 \leqq f(x) \leqq 1$ for each $x \in S$, so that there exists a continuous extension φ of f to the domain $\beta(S)$. Letting the bar denote closures in the space $\beta(S)$, we have $\varphi(x) = 0$ for each $x \in \bar{A}_1$ and $\varphi(x) = 1$ for each $x \in \bar{A}_2$, so that indeed $\bar{A}_1\bar{A}_2 = 0$. Conversely, let the space B have properties (1), (2), (3), (4′). There exists a continuous mapping h of the space $\beta(S)$ into the space B such that $h(x) = x$ for each $x \in S$. It is sufficient to prove that the mapping h is one-to-one. Suppose the contrary. Then there exist two points $a \in \beta(S)$, $b \in \beta(S)$ such that $a \neq b$, $h(a) = h(b)$. There exists a continuous function f in the domain $\beta(S)$ such that $f(a) = 0, f(b) = 1$. Let A_1 denote the set of all points $x \in S$ such that $f(x) \leqq \frac{1}{3}$; let A_2 denote the set of all points $x \in S$ such that $f(x) \geqq \frac{2}{3}$. It is easy to see that A_1 and A_2 are two completely separated subsets of S so that $\bar{A}_1\bar{A}_2 = 0$ where the bar designates closures in the space B. Since $h(a) = h(b)$, we shall have a contradiction if we shall prove that $h(a) \in \bar{A}_1$, $h(b) \in \bar{A}_2$. Let U be any neighborhood of $h(a)$ in the space B. Then $h^{-1}(U)$ is a neighborhood of a in the space $\beta(S)$. Since $f(a) = 0$ and since S is dense in $\beta(S)$, it is easy to see that $h^{-1}(U) \cdot A_1 \neq 0$, whence $U \cdot A_1 \neq 0$. Since U was an arbitrary neighborhood of $h(a)$ in the space B, we have indeed $h(a) \in \bar{A}_1$ and similarly we prove that $h(b) \in \bar{A}_2$.

In the particular case when *S is a normal Riesz space*, it follows from the result just proved that $\beta(S)$ may characterized by the properties (1), (2), (3) and (5) where (5) means the following: *If F_1 and F_2 are two closed subsets of S without common points, then the closures of F_1 and F_2 in the space $\beta(S)$ have no common points.* Conversely, *if there exists a space B having properties* (1), (2), (3) *and* (5), *then S is normal and $B = \beta(S)$.* Indeed, it is easy to see that property (5) is stronger than property (4′) so that $B = \beta(S)$. If F_1 and F_2 are two closed subsets of S and $F_1F_2 = 0$, then $\bar{F}_1\bar{F}_2 = 0$, the bar indicating closures in B. Since B is a bicompact Hausdorff space, it is normal, so that there exists a continuous function φ in the domain $\beta(S)$ such that $\varphi(x) = 0$ for each $x \in \bar{F}_1$ and $\varphi(x) = 1$ for each $x \in \bar{F}_2$. Hence it follows that S is normal.

Let S be a completely regular space. Let T be a closed subset of S; let $\bar{T}$ denote the closure of T in the space $\beta(S)$. Then *we have* $\bar{T} = \beta(T)$ (i.e. $\bar{T}$ possesses the properties (1)-(4) of $\beta(T)$) *if and only if every bounded*[7] *continuous function in the domain T admits of a continuous extension to the domain S.* Suppose first that $\bar{T} = \beta(T)$ and let f be a continuous function in the domain T such that e.g. $0 \leqq f(x) \leqq 1$ for each $x \in T$. Since $\bar{T} = \beta(T)$, there exists a continuous extension g of f to the domain $\bar{T}$; of course $0 \leqq g(x) \leqq 1$ for each $x \in \bar{T}$. Since $\beta(S)$ is a bicompact Hausdorff space, it is normal; since $\bar{T}$ is closed in

$\beta(S)$, there exists a continuous extension φ of g to the domain $\beta(S)$. Hence f may be continuously extended to the domain $\beta(S)$ and therefore also to the domain $S \subset \beta(S)$. Conversely suppose that every bounded continuous function in the domain T may be continuously extended to the domain S. Of course $\overline{T}$ has always properties (1)-(3) (relatively to T); therefore to prove that $\overline{T} = \beta(T)$ it is sufficient to prove that $\overline{T}$ has property (4′) (again relatively to T). Hence suppose that $A_1 \subset T$ and $A_2 \subset T$ are completely separated in the space T. Then there exists a continuous function f in the domain T such that $f(x) = 0$ for each $x \in A_1$, $f(x) = 1$ for each $x \in A_2$ and $0 \leqq f(x) \leqq 1$ for each $x \in T$. There exists a continuous extension φ of f to the domain S, whence it readily follows that A_1 and A_2 are completely separated in the space S. Since $\beta(S)$ has property (4′) (relatively to S), we have $\bar{A}_1\bar{A}_2 = 0$, the bar indicating closures in the space $\beta(S)$. But of course $\bar{A}_1$ and $\bar{A}_2$ are closures of A_1 and A_2 in the space $\overline{T}$, so that $\overline{T}$ has indeed property (4′) relatively to T.

The theorem just proved has the following consequence: *If S is a normal Riesz space, then* $\overline{T} = \beta(T)$ (the bar indicating closure in $\beta(S)$) *for every closed subset T of S. If the completely regular space S is not normal, then there exists a closed subset T of S such that $\overline{T} \neq \beta(T)$.*

If Φ is a family of neighborhoods of a point x of a topological space S, then we say that Φ is *complete* if, given an arbitrary neighborhood G of x, there exists a neighborhood U of x such that both $U \in \Phi$ and $U \subset G$. The least cardinal number of a complete family of neighborhoods of x is called the *character*[13] of x (in the space S) and is denoted by $\chi(x) = \chi_S(x)$. If $T \subset S$ and $x \in T$, it is easy to see that

$$\chi_T(x) \leqq \chi_S(x).$$

Let S be a completely regular space. Then for every point $a \in S$ we have

$$\chi_S(a) = \chi_{\beta(S)}(a).$$

Let Φ be a complete family of neighborhoods of a in the space S whose cardinal number is equal to $\chi_S(a)$. It is sufficient to construct a complete family Ψ of neighborhoods of a in the space $\beta(S)$ such that the cardinal number of Ψ does not exceed $\chi_S(a)$. The family Ψ will be constructed as a transform of the family Φ, each $U \in \Phi$ determining a $\tau(U) \in \Psi$, in the following way,

$$\tau(U) = \beta(S) - \overline{S - U}$$

(the bar indicating closures in the space $\beta(S)$). Of course Ψ is a family of neighborhoods of a in the space $\beta(S)$ and the cardinal number of Ψ does not exceed $\chi_S(a)$. Hence we have only to prove that, given a neighborhood G of a in the space $\beta(S)$, there exists a $U \in \Phi$ such that $\tau(U) \subset G$. There exists a continuous function f in the domain $\beta(S)$ such that $f(a) = 0$ and $f(x) = 1$ for each $x \in \beta(S) - G$. Let H denote the set of all points $x \in S$ such that $f(x) < \frac{1}{2}$. Then H is a neighborhood of a in the space S, so that there exists a $U \in \Phi$ such that $U \subset H$.

[13] AU, p. 2.

It remains to prove that $\tau(U) \subset G$. Supposing the contrary, there exists a point $b \in \tau(U) - G$. Since $b \in \beta(S) - G$, we have $f(b) = 1$. Let V be an arbitrary neighborhood of b in the space $\beta(S)$. Since $f(b) = 1$ and since S is dense in $\beta(S)$, there exists a point $c \in SV$ such that $f(c) > \frac{1}{2}$. Since $U \subset H$, we cannot have $c \in U$. Therefore $c \in S - U$ so that $(S - U) V \neq 0$. Since V was an arbitrary neighborhood of b in the space $\beta(S)$, we have $b \in \overline{S - U} = \beta(S) - \tau(U)$, which is a contradiction.

Let S be a completely regular space. Let $A \subset \beta(S) - S (A \neq 0)$ be both closed and a G_δ in $\beta(S)$. Then the cardinal number of A is $\geqq 2^{\aleph_0}$. Since A is both closed and a G_δ in the normal space $\beta(S)$, there exists a continuous function f in the domain $\beta(S)$ such that $f(x) = 0$ for each $x \in A$ and $f(x) > 0$ for each $x \in \beta(S) - A$. The set of all points $x \in \beta(S)$ such that $f(x) < n^{-1} (n = 1, 2, 3, \cdots)$ is open and not vacuous. Since S is dense in $\beta(S)$, there exists a point $a_n \in S$ such that $f(a_n) < n^{-1}$. Since $AS = 0$, we have $f(a_n) > 0$. It is evident that the points a_n may be chosen is such a manner that $f(a_{n+1}) < f(a_n)$. Let us arrange the rational numbers of the interval $0 < t < 1$ in a simple sequence $\{r_n\}$. There exists a continuous function φ in the domain $0 < t < \infty$ such that $0 < \varphi(t) < 1$ and $\varphi[f(a_n)] = r_n (n = 1, 2, 3, \cdots)$. Since $f(x) > 0$ for each $x \in S$, we obtain a bounded continuous function g in the domain S such that $g(x) = \varphi[f(x)]$ for each $x \in S$. There exists a continuous extension h of g to the domain $\beta(S)$. Choose a real number $\alpha, 0 \leqq \alpha \leqq 1$. There exists a sequence $i_1 < i_2 < i_3 < \cdots$ such that $r_{i_n} \to \alpha$ for $n \to \infty$. Let M_n designate the set of points a_{i_n}, $a_{i_{n+1}}, a_{i_{n+2}}, \cdots$ so that $M_n \subset S$, $M_n \supset M_{n+1}$, $M_n \neq 0$. Since the space $\beta(S)$ is bicompact, there exists a point $b \in \prod \bar{M}_n$. Since the functions f and h are continuous, we have $f(\bar{M}_n) \subset \overline{f(M_n)}$, $h(\bar{M}_n) \subset \overline{h(M_n)} = \overline{g(M_n)}$, whence $f(b) \in \prod \overline{f(M_n)}$, $h(b) \in \prod \overline{g(M_n)}$. Since $f(a_{i_n}) \to 0$, $g(a_{i_n}) \to \alpha$ for $n \to \infty$, we easily see that $f(b) = 0$, $h(b) = \alpha$. Since $f(b) = 0$, we have $b \in A$. Therefore, for each α such that $0 \leqq \alpha \leqq 1$, the set A contains a point b such that $h(b) = \alpha$. Hence the cardinal number of A is at least $2^{\aleph_0}$.

Let S_1 and S_2 be two completely regular spaces satisfying the first countability axiom. Let the spaces $\beta(S_1)$ and $\beta(S_2)$ be homeomorphic. Then the spaces S_1 and S_2 are homeomorphic. We may assume that $\beta(S_1) = \beta(S_2)$. According to the preceding theorem no point $x \in \beta(S_1) - S_1$ is a G_δ in $\beta(S_1)$. But every point $x \in S_2$ satisfies the first countability axiom relatively to S_2 and, therefore, after the theorem last but one, relatively to $\beta(S_2)$ as well and hence x is a G_δ in $\beta(S_2) = \beta(S_1)$. Therefore $S_2 \subset S_1$ and similarly $S_1 \subset S_2$, so that $S_1 = S_2$.

Let I denote an infinite countable isolated space (e.g. the space of all natural numbers). It is an important problem to determine the cardinal number $\mathfrak{m}$ of $\beta(I)$. All I know about it is that

$$2^{\aleph_0} \leqq \mathfrak{m} \leqq 2^{2^{\aleph_0}}.$$

It is easily seen that each point of I is an isolated point of $\beta(I)$ so that the set I is open in $\beta(I)$. Since I is countable, it is an F_σ in $\beta(I)$. Hence $\beta(I) - I$ is both closed and a G_δ in $\beta(I)$ so that the cardinal number of $\beta(I) - I$ is $\geqq 2^{\aleph_0}$.

On the other hand, since the set I is dense in the Hausdorff space $\beta(I)$, it is easy to see that a point $x \in \beta(I)$ is uniquely determined knowing the family of all sets $A \subset I$ such that $x \in \bar{A}$, so that the cardinal number of $\beta(I)$ is at most equal to the cardinal number $2^{2^{\aleph_0}}$ of all families of subsets of I.

A topological space S is called *compact* if, given any infinite subset A of S, there exists a point $x \in S$ such that $x \in \overline{A - x}$.

Let the normal Riesz space S be not compact. Then the cardinal number of $\beta(S) - S$ is at least equal to the cardinal number of $\beta(I)$ (hence at least equal to $2^{\aleph_0}$). Since S is not compact, it is well known that S contains a closed subset F homeomorphic with I. Since S is normal, we have $\beta(I) = \bar{I} \subset \beta(S)$, so that $\beta(I) - I \subset \beta(S) - S$. But the sets $\beta(I) - I$ and $\beta(I)$ have the same cardinal number.

I do not know whether this theorem remains true if we replace normality by complete regularity. It may be shown that the assumption of normality may be replaced by the following weaker assumption[14]: If F_1 and F_2 are two closed subsets of S such that F_1 *is countable* and $F_1F_2 = 0$, there exist two open sets G_1 and G_2 such that $G_1 \supset F_1$, $G_2 \supset F_2$, $G_1G_2 = 0$.

If the space S is compact, then the set $\beta(S) - S$ may consist of a single point. Let S be the set of all ordinal numbers $<\omega_1$, ω_1 being the first uncountable ordinal number. Let S_0 be the set of all ordinal numbers $\leqq \omega_1$. The topology of S and S_0 is the usual topology of an ordered set, an open base being given by the family of all open intervals. It is well known that S is a compact normal Riesz space and that S_0 is a bicompact Hausdorff space. We shall prove that $S_0 = \beta(S)$. Since it is evident that S_0 possesses properties (1)–(3) of $\beta(S)$, it is sufficient to prove that a continuous function f in the domain S admits of a continuous extension to the domain S_0. This is an easy consequence of the following theorem. *If f is a continuous function in the domain S, then there exists a point $\xi \in S$ such that f is constant for $x \geqq \xi$.* It is sufficient to prove that, given a number $\varepsilon > 0$, there exists a point $\xi(\varepsilon) \in S$ such that $|f(x) - f(y)| < \varepsilon$ for $x \in S, y \in S, x > \xi(\varepsilon), y > \xi(\varepsilon)$. Supposing the contrary, there would exist in S two sequences $\{a_n\}$ and $\{b_n\}$ such that $a_n < b_n < a_{n+1}$ and $|f(a_n) - f(b_n)| \geqq \varepsilon$. But this is impossible, because f would then be discontinuous at α, α being the first ordinal number greater than each a_n.

We say that $x \in S$ is a *κ-point*[15], if there exists a sequence $\{x_n\} \subset S - (x)$ such that $\lim x_n = x$, i.e. that, given any neighborhood U of x, we have $x_n \in U$ except for a finite number of subscripts n. Alexandroff and Urysohn raised the question[16] whether there exists a bicompact Hausdorff space which is dense in itself and which contains no κ-point. We shall prove that the space $\beta(I) - I$ has this property. Supposing the contrary, there exists a point $c \in \beta(I) - I$ and a sequence $\{a_n\} \subset \beta(I) - I - (c)$ such that $\lim a_n = c$. We may suppose that the points a_n are all distinct from one another. Let A_n be the set of the points

[14] AU, p. 58.
[15] AU, p. 53.
[16] AU, p. 54.

$a_n, a_{n+1}, a_{n+2} \cdots$ together with the point c. It is easy to see that A_n is a closed subset of $\beta(I)$. We shall construct successively open subsets U_n of the space $\beta(I)$ as follows. U_1 contains the point a_1, but $\overline{U}_1 A_2 = 0$. If, for a certain value of n, we have already constructed the set U_n so that $\overline{U}_n \cdot A_{n+1} = 0$, let U_{n+1} be an open subset containing a_{n+1}, but such that $\overline{U}_{n+1} \cdot \overline{U}_i = 0$ for $1 \leqq i \leqq n$ and $\overline{U}_{n+1} \cdot A_{n+2} = 0$. It is easy to see that the successive construction of the sequence $\{U_n\}$ may be carried through. Now put $\Phi = I \cdot \sum U_{2n-1}$, $\Psi = I \cdot \sum U_{2n}$. Then $\Phi\Psi = 0$ and the sets Φ and Ψ are of course closed in I, since I is an isolated space. Since I is normal, we must have $\bar{\Phi}\bar{\Psi} = 0$, the bars indicating closures in $\beta(I)$. On the other hand, since I is dense in $\beta(I)$ and U_n is open in $\beta(I)$, it is easy to see that $\overline{IU_n} = \overline{U}_n$, so that $a_n \in \overline{IU_n}$, whence we easily get the contradiction $c \in \bar{\Phi}\bar{\Psi}$.

III

We shall say that the space S is *topologically complete* if there exists a bicompact Hausdorff space $B \supset S$ such that S is a G_δ in B. Of course S is then completely regular. *A G_δ in a topologically complete space is a topologically complete space. A closed subset of a topolologically complete space is a topologically complete space.*

A topological space S is topologically complete if and only if it is completely regular and a G_δ in $\beta(S)$. If S is a G_δ in $\beta(S)$, then it is topologically complete, since $\beta(S)$ is a bicompact Hausdorff space. Conversely suppose that S is topologically complete. Then there exists a bicompact Hausdorff space $B \supset S$ such that S is a G_δ in B. Let B_0 be the closure of S in the space B. Then B_0 is a bicompact Hausdorff space and S is dense in B_0 and a G_δ in B_0. We know that there exists a continuous mapping h of $\beta(S)$ into B_0 such that $h^{-1}(S) = S$. Since S is a G_δ in B_0, it is easy to see that $h^{-1}(S) = S$ is a G_δ in $\beta(S)$.

Let T be a completely regular[17] *space. Let $S \subset T$ be a topologically complete space. Then S is a G_δ in the closure of S in the space T.* Let S_0 be the closure of S in the space $\beta(T)$. It is sufficient to prove that S is a G_δ in S_0. Since S_0 is a bicompact Hausdorff space and since S is dense in S_0, there exists a continuous mapping h of $\beta(S)$ into S_0 such that $h[\beta(S) - S] = S_0 - S$. Since S is topologically complete, it is a G_δ in $\beta(S)$, so that $\beta(S) - S$ is an F_σ in $\beta(S)$. Hence there exist closed subsets F_n of $\beta(S)$ such that $\sum F_n = \beta(S) - S$, whence $S_0 - S = \sum h(F_n)$. Every F_n is a bicompact space, so that every $h(F_n)$ is a bicompact space. Since $h(F_n)$ is a bicompact subset of the Hausdorff space S_0, it is closed in S_0, so that $S_0 - S$ is an F_σ in S_0 and finally S is a G_δ in S_0.

Let T be a topologically complete space. Let $S \subset T$. Then S is a topologically complete space if and only if it is the intersection of a closed subset of T and a G_δ in T. If $S = FH$, where F is closed in T and H is a G_δ in T, then F is a topologically complete space and S is a G_δ in F, so that S is a topologically complete space. Conversely let S be topologically complete. Then S is a G_δ in the closure $\bar{S}$ of S in T, so that $S = \bar{S}H$, H being a G_δ in T.

---

[17] I do not know whether this assumption is necessary.

Let $S \neq 0$ be a topologically complete space[18]. *Let $\{G_n\}$ be a sequence of open and dense subsets of S. Let $H = \prod G_n$. Then $H \neq 0$ and, moreover, H is dense in S.* There exists a regular compact (as a matter of fact, bicompact) space $K \supset S$ such that S is a G_δ in K. We may suppose that $\bar{S} = K$, the bar denoting closure in K. The sets G_n being open in S, there exist sets Γ_n open in K and such that $G_n = S \cdot \Gamma_n$. Since S is a G_δ in K, there exist sets Δ_n open in K and such that $S = \prod \Delta_n$. Since S is dense in K and G_n are dense in S, the sets G_n are dense in K. Choose an arbitrary point $a_0 \in S$ and an arbitrary neighborhood V of a_0 in the space S. All we have to prove is that $HV \neq 0$. There exists a neighborhood U_0 of a_0 in the space K such that $V = SU_0$. Since the set G_1 is dense in K, there exists a point $a_1 \in G_1 U_0 = S \cdot \Gamma_1 U_0 \subset \Delta_1 \Gamma_1 U_0$. Hence $\Delta_1 \Gamma_1 U_0$ is a neighborhood of a_1 in the space K. Since K is regular, there exists a neighborhood U_1 of a_1 (in the space K) such that $\bar{U}_1 \subset \Delta_1 \Gamma_1 U_0$. Generally, let there be given for a certain value of n a point $a_n \in G_n$ and its neighborhood U_n (in the space K) such that $\bar{U}_n \subset \Delta_n \Gamma_n U_{n-1}$. Then $a_n \in G_n \subset S$ and SU_n is a neighborhood of a_n in the space S; since G_{n+1} is dense in S, there exists a point $a_{n+1} \in G_{n+1} U_n = S \cdot \Gamma_{n+1} U_n \subset \Delta_{n+1} \Gamma_{n+1} U_n$. Hence $\Delta_{n+1} \Gamma_{n+1} U_n$ is a neighborhood of a_{n+1} in the regular space K, so that there exists a neighborhood U_{n+1} of a_{n+1} (in the space K) such that $\bar{U}_{n+1} \subset \Delta_{n+1} \Gamma_{n+1} U_n$. Thus we construct a sequence $\{a_n\}$ of points and a sequence $\{U_n\}$ of open sets so that $a_n \in G_n U_n$, $\bar{U}_{n+1} \subset \Delta_{n+1} \Gamma_{n+1} U_n$. Since $a_n \in U_n$, we have $U_n \neq 0$. Since K is compact and $\bar{U}_{n+1} \subset U_n$, there exists a point $b \in \prod U_n = \prod \bar{U}_n$. Since $\bar{U}_{n+1} \subset \Delta_{n+1} \Gamma_{n+1} U_n$, we have $b \in \prod \Delta_n$. $\prod \Gamma_n = S \cdot \prod \Gamma_n = \prod G_n = H$. Moreover $b \in U_0$, so that $b \in HU_0 = HV$.

Let S be a metric space. A *Cauchy sequence* in S is a sequence $\{x_n\} \subset S$ such that, given a number $\varepsilon > 0$, there exists a number p such that the distance of x_m and x_n is less than ε, whenever both m and n are greater than p. A metric space S is called *metrically complete* if, given any Cauchy sequence $\{x_n\}$ in S, there exists a point $x \in S$ such that $\lim x_n = x$. A topological space is called *completely metrizable*, if it is homeomorphic with a metrically complete space.

We next prove our principal theorem: *A metrizable space S is topologically complete if and only if it is completely metrizable.*

Let S be a metrically complete space and let ρ be its distance function. We may suppose that $\rho(x, y) \leqq 1$ for every pair of points, since otherwise we may replace ρ by ρ_1, putting $\rho_1(x, y) = \rho(x, y)$ if $\rho(x, y) \leqq 1$, $\rho_1(x, y) = 1$ if $\rho(x, y) > 1$ Since S is metric, it is completely regular, so that $\beta(S)$ exists. For any given $a \in S$, $\rho(a, x)$ is a bounded continuous function in the domain S so that there exists a continuous function $\varphi_a(x)$ in the domain $\beta(S)$ such that $\varphi_a(x) = \rho(a, x)$ for each $x \in S$. If $a \in S$, $b \in S$, then the set $T(a, b)$ of all points $x \in \beta(S)$ such that $\varphi_a(x) + \varphi_b(x) \geqq \rho(a, b)$ is closed in $\beta(S)$ and contains S. Since S is dense in $\beta(S)$, we must have $T(a, b) = \beta(S)$, i.e. $\varphi_a(x) + \varphi_b(x) \geqq \rho(a, b)$ for each $x \in \beta(S)$.

---

[18] It is evident from the proof that it is possible to replace this by the weaker assumption that S is a G_δ in some regular compact space.

For $a \in S$ and $n = 1, 2, 3, \cdots$ let $\Gamma(a, n)$ be the set of all points $x \in \beta(S)$ such that $\varphi_a(x) < n^{-1}$. Since the function $\varphi_a(x)$ is continuous, $\Gamma(a, n)$ is an open subset of $\beta(S)$. Therefore

$$G_n = \sum_{a \in S} \Gamma(a, n)$$

is an open set. We shall prove that $S = \prod G_n$, so that the set S is a G_δ in $\beta(S)$ and thus topologically complete. Evidently $\prod G_n \supset S$. Conversely let $b \in \prod G_n$. We have to prove that $b \in S$. According to the definition of G_n, there exist points $a_n \in S$ such that $\varphi_{a_n}(b) < n^{-1}$. Therefore

$$\rho(a_n, a_m) \leqq \varphi_{a_n}(b) + \varphi_{a_m}(b) < \frac{1}{n} + \frac{1}{m},$$

so that $\{a_n\}$ is a Cauchy sequence in S. Since S is metrically complete, there exists a point $a \in S$ such that $a = \lim a_n$. It is sufficient to prove that $a = b$. Suppose that $a \neq b$. Since $\beta(S)$ is a Hausdorff space, there exist two open subsets U and V of $\beta(S)$ such that $a \in U$, $b \in V$, $UV = 0$. Since US is a neighborhood of a in the metric space S, there exists an integer $n > 0$ such that U contains every point $x \in S$ such that $\rho(a, x) < 2 \cdot n^{-1}$. This can be written in the form $SW \subset U$, W being the set of all points $x \in \beta(S)$ such that $\varphi_a(x) < 2 \cdot n^{-1}$. Since φ_a is continuous, W is an open subset of $\beta(S)$. Since S is dense in $\beta(S)$ and U, V and W are open in $\beta(S)$, we have $W \subset \overline{W} = \overline{SW} \subset \overline{U} \subset \beta(S) - V$, or $WV = 0$. Hence for each $x \in V$ we have $\varphi_a(x) \geqq 2 \cdot n^{-1}$; in particular $\varphi_a(b) \geqq 2 \cdot n^{-1}$. Since $\rho(a_n, a_m) < n^{-1} + m^{-1}$ and $\lim a_n = a$, we have $\rho(a, a_n) \leqq n^{-1}$. Hence for each $x \in S$ we have $\rho(a, x) \leqq \rho(a, a_n) + \rho(a_n, x) \leqq n^{-1} + \rho(a_n, x)$, whence it easily follows that for each $x \in \beta(S)$ we have $\varphi_a(x) \leqq \varphi_{a_n}(x) + n^{-1}$, in particular $\varphi_a(b) \leqq \varphi_{a_n}(b) + n^{-1} < n^{-1} + n^{-1} = 2 \cdot n^{-1}$, which is a contradiction.

Now suppose that the metric space S is topologically complete. Let ρ denote the distance function of S; again, we shall suppose that $\rho(x, y) \leqq 1$ for every couple of points. Since S is topologically complete, there exists a sequence $\{F_n\}$ of closed subsets of $\beta(S)$ such that $\beta(S) - S = \sum F_n$. If $S = \beta(S)$, then S is a bicompact metric space, and then it is well known that S is metrically complete. Hence let us suppose that $S \neq \beta(S)$; we may then assume that $F_n \neq 0$ for every n. Given any point $a \in S$, $\rho(a, x)$ is a bounded continuous function in the domain S, which admits of a continuous extension φ_a to the domain $\beta(S)$. If the point $b \in \beta(S)$ is different from a, then there exist open subsets U and V of $\beta(S)$ such that $a \in U$, $b \in V$, $UV = 0$. Since SU is a neighborhood of a in the metric space S, there exists a number $\varepsilon > 0$ such that U contains every point $x \in S$ such that $\rho(a, x) < \varepsilon$. Since S is dense in $\beta(S)$, it easily follows that $\overline{U}$ contains every point $x \in \beta(S)$ such that $\varphi_a(x) < \varepsilon$. Since $U \subset \beta(S) - V = \overline{\beta(S) - V}$, we have $\overline{U} \subset \beta(S) - V$ so that $b \in \beta(S) - \overline{U}$, whence $\varphi_a(b) \geqq \varepsilon$. Thus we proved that $\varphi_a(b) > 0$ for every $b \in \beta(S)$ except for $b = a$. Since the set $F_n \neq 0$ is closed in the bicompact space $\beta(S)$, it is easy to see that the function $\varphi_a(x)$, x running over F_n, admits of a minimum value $\sigma(a, F_n)$. Since $a \in S$, $F_n S = 0$, we have $\sigma(a, F_n) > 0$.

If $a \epsilon S$, $b \epsilon S$, then we have $\rho(a, x) \leqq \rho(a, b) + \rho(b, x)$ for every $x \epsilon S$, whence $\varphi_a(x) \leqq \rho(a, b) + \varphi_a(x)$ for every $x \epsilon \beta(S)$. Therefore $\sigma(a, F_n) \leqq \rho(a, b) + \sigma(b, F_n)$, and similarly $\sigma(b, F_n) \leqq \rho(a, b) + \sigma(a, F_n)$. Hence

$$| \sigma(a, F_n) - \sigma(b, F_n) | \leqq \rho(a, b).$$

Now let us put for $x \epsilon S$, $y \epsilon S$

$$f_n(x, y) = \rho(x, y) + \sigma(x, F_n) + \sigma(y, F_n),$$

$$g_n(x, y) = \frac{\rho(x, y)}{f_n(x. y)},$$

$$\rho_0(x, y) = \rho(x, y) + \sum_1^\infty 2^{-n} \cdot g_n(x, y).$$

Since $\rho(x, y) \geqq 0$, $\sigma(x, F_n) > 0$, $\sigma(y, F_n) > 0$, we have $f_n(x, y) > 0$. Hence $g_n(x, y)$ exists and $0 \leqq g_n(x, y) \leqq 1$, so that the series $\sum 2^{-n} \cdot g_n(x, y)$ is convergent. It is evident that $\rho_0(x, y) = \rho_0(y, x)$ and that $\rho_0(x, x) = 0$, whereas $\rho_0(x, y) > 0$ if $x \neq y$. Next we shall prove that $\rho_0(x, z) \leqq \rho_0(x, y) + \rho_0(y, z)$ for $x \epsilon S$, $y \epsilon S$, $z \epsilon S$. Since

$$\frac{t_1}{c + t_1} \leqq \frac{t_2}{c + t_2} \quad \text{for } c > 0,\ 0 \leqq t_1 \leqq t_2$$

and since $0 \leqq \rho(x, z) \leqq \rho(x, y) + \rho(y, z)$, we have

$$g_n(x, z) \leqq \frac{\rho(x, y) + \rho(y, z)}{\rho(x, y) + \rho(y, z) + \sigma(x, F_n) + \sigma(z, F_n)}.$$

Since

$$\sigma(y, F_n) \leqq \rho(x, y) + \sigma(x, F_n),$$

$$\sigma(y, F_n) \geqq \rho(y, z) + \sigma(z, F_n),$$

we have

$$\rho(x, y) + \rho(y, z) + \sigma(x, F_n) + \sigma(z, F_n) \geqq \begin{cases} \rho(x, y) + \sigma(x, F_n) + \sigma(y, F_n), \\ \rho(y, z) + \sigma(y, F_n) + \sigma(z, F_n), \end{cases}$$

whence

$$g_n(x, z) \leqq g_n(x, y) + g_n(y, z),$$

so that indeed

$$\rho_0(x, z) \leqq \rho_0(x, y) + \rho_0(y, z).$$

Hence ρ_0 has all the properties of a distance function. Next we prove that ρ and ρ_0 are equivalent metrics in S, i.e. that for $x \epsilon S$ and $\{x_n\} \subset S$ we have

$$\lim \rho(x_n, x) = 0 \text{ if and only if } \lim \rho_0(x_n, x) = 0.$$

If $\lim \rho_0(x_n, x) = 0$, then $\lim \rho(x_n, x) = 0$, since $0 \leqq \rho(x_n, x) \leqq \rho_0(x_n, x)$. Conversely suppose that $\lim \rho(x_n, x) = 0$. Choose a number $\varepsilon > 0$ and an integer $k > 0$ such that $2^{-k+1} < \varepsilon$. Then we have for all values of n

$$\sum_{i=k+1}^{\infty} 2^{-i} g_i(x_n, x) \leqq \sum_{i=k+1}^{\infty} 2^{-i} = 2^{-k} < \tfrac{1}{2}\varepsilon,$$

whence

$$\begin{aligned} \rho_0(x_n, x) &< \rho(x_n, x) + \sum_{i=1}^{k} 2^{-i} g_i(x_n, x) + \tfrac{1}{2}\varepsilon \\ &\leqq \rho(x_n, x) + \sum_{i=1}^{k} 2^{-i} \frac{\rho(x_n, x)}{\rho(x_n, x) + \sigma(x, F_i)} + \tfrac{1}{2}\varepsilon. \end{aligned}$$

Since $\lim \rho(x_n, x) = 0$, we must have

$$\lim_{n\to\infty} \sum_{i=1}^{k} 2^{-i} \frac{\rho(x_n, x)}{\rho(x_n, x) + \sigma(x, F_i)} = 0,$$

so that there exists an integer p such that for $n > p$ we have

$$0 \leqq \sum_{i=1}^{k} 2^{-i} \frac{\rho(x_n, x)}{\rho(x_n, x) + \sigma(x, F_i)} < \tfrac{1}{2}\varepsilon.$$

Therefore

$$\rho_0(x_n, x) < \rho(x_n, x) + \varepsilon$$

for every $n > p$. Since $\lim \rho(x_n, x) = 0$ and the number $\varepsilon > 0$ was arbitrary, we have indeed $\lim \rho_0(x_n, x) = 0$. Thus we proved that ρ and ρ_0 are equivalent metrics in S, i.e. that the metric spaces $S = (S, \rho)$ and (S, ρ_0) are homeomorphic.

It remains to be shown that the metric space (S, ρ_0) is metrically complete. Hence suppose that $\{x_n\}$ is a Cauchy sequence in (S, ρ_0). We have to prove that there exists a point $x \in S$ such that $\lim \rho_0(x_n, x) = 0$, or, what we already know to be equivalent, that $\lim \rho(x_n, x) = 0$. Since the space $\beta(S)$ is bicompact, it is easy to see that there exists a point $x \in \beta(S)$ such that, given any neighborhood U of x (in the space $\beta(S)$), we have $x_n \in U$ for an infinite number of values of n. It is sufficient to prove that $x \in S$, for then, since $\{x_n\}$ is a Cauchy sequence, it is easy to show that $\lim \rho(x_n, x) = 0$. Suppose, on the contrary, that the point x belongs to the set $\beta(S) - S = \sum F_n$. Hence there exists an integer $k > 0$ such that $x \in F_k$.

We shall prove that $\sigma(x_n, F_k) \to 0$ for $n \to \infty$. Choose a number $\varepsilon > 0$. There exists an integer $p > 0$ such that for $n > p$, $m > p$ we have $\rho(x_n, x_m) \leqq \rho_0(x_n, x_m) < \varepsilon$. Let n be greater than p. The number $\sigma(x_n, F_k)$ is the minimum value of $\varphi_{x_n}(y)$ for $y \in F_k$. Since $x \in F_k$, we must have $0 < \sigma(x_n, F_k) \leqq \varphi_{x_n}(x)$. There exists a neighborhood Ω_n of x in $\beta(S)$ such that $|\varphi_{x_n}(z) - \varphi_{x_n}(x)| < \varepsilon$ for every $z \in \Omega_n$. There exists an integer $m_n > p$ such that $x_{m_n} \in \Omega_n$, whence $|\varphi_{x_n}(x_{m_n}) - \varphi_{x_n}(x)| < \varepsilon$, i.e. $|\rho(x_n, x_{m_n}) - \varphi_{x_n}(x)| < \varepsilon$. Since $n > p$, $m_n > p$, we must have $\rho(x_n, x_{m_n}) < \varepsilon$, whence $\varphi_{x_n}(x) < 2\varepsilon$. Therefore $0 < \sigma(x_n, F_k) < 2\varepsilon$ for $n > p$, so that indeed $\sigma(x_n, F_k) \to 0$ for $n \to \infty$.

Since $\{x_n\}$ is a Cauchy sequence in (S, ρ_0), there exists an integer p such that $\rho_0(x_n, x_p) < 2^{-k-2}$ for each $n > p$. But

$$\rho_0(x_n, x_p) \geqq 2^{-k} g_k(x_n, x_p) = 2^{-k} \frac{\rho(x_n, x_p)}{\rho(x_n, x_p) + \sigma(x_n, F_k) + \sigma(x_p, F_k)}.$$

Since

$$\sigma(x_p, F_k) \leqq \rho(x_n, x_p) + \sigma(x_n, F_k),$$

it follows that

$$\rho_0(x_n, x_p) \geqq 2^{-k-1} \frac{\rho(x_n, x_p)}{\rho(x_n, x_p) + \sigma(x_n, F_k)} \geqq 0,$$

so that for every $n > p$ we have

$$0 \leqq \frac{\rho(x_n, x_p)}{\rho(x_n, x_p) + \sigma(x_n, F_k)} < \tfrac{1}{2},$$

whence $\rho(x_n, x_p) < \sigma(x_n, F_k)$. But $\sigma(x_n, F_k) \to 0$ f[illegible] ∞. Therefore $\rho(x_n, x_p) \to 0$ for $n \to \infty$. Hence there exists an integer q [illegible] h that for every $n > q$ we have $\rho(x_n, x_p) < \frac{1}{2}\, \varphi_{x_p}(x)$. [Since $x_p \in S$, $x \in \beta(S) - S$, we know that $\varphi_{x_p}(x) > 0$.] There exists a neighborhood U of x in the space $\beta(S)$ such that $\varphi_{x_p}(z) > \frac{1}{2}\varphi_{x_p}(x)$ for any $z \in U$. There exists an integer $n > q$ such that $x_n \in U$, whence $\rho(x_n, x_p) = \varphi_{x_p}(x_n) > \frac{1}{2}\varphi_{x_p}(x)$, which is a contradiction.

IV

Let S be a completely regular space. Let $\lambda(S)$ be the set of all points $x \in \beta(S)$ such that x possesses a neighborhood U(in the space $\beta(S)$) such that $S \cdot \overline{U}$ is a normal space. [$\overline{U}$ is the closure of U in $\beta(S)$]. It is easy to see that $\lambda(S)$ is an open subset of $\beta(S)$.

Let F_1 and F_2 be two closed subsets of a completely regular space S such that $F_1 F_2 = 0$. Then

$$\bar{F}_1 \cdot \bar{F}_2 \cdot \lambda(S) = 0,$$

the bars indicating closures in $\beta(S)$. Supposing the contrary, there exists a point $a \in \bar{F}_1 \cdot \bar{F}_2 \cdot \lambda(S)$. Since $a \in \lambda(S)$, there exists a neighborhood U of a (in the space $\beta(S)$) such that $S \cdot \overline{U}$ is a normal space. There exists a neighborhood V of a such that $\bar{V} \subset U$. Put

$$\Phi_1 = \bar{V} \cdot F_1, \; \Phi_2 = \overline{U} \cdot F_2 + S(\overline{U} - U).$$

Then Φ_1 and Φ_2 are two closed subsets of $S\overline{U}$ such that $\Phi_1 \Phi_2 = 0$. Moreover, it is easy to see that $a \in \bar{\Phi}_1 \cdot \bar{\Phi}_2$. Since $S\overline{U}$ is a normal space, there exists a bounded continuous function f in the domain $S\overline{U}$ such that $f(x) = 0$ for each $x \in \Phi_1$ and $f(x) = 1$ for each $x \in \Phi_2$. For $x \in S$ put (i) $g(x) = f(x)$ if $x \in SU$, (ii) $g(x) = 1$ if $x \in S - U$. Then it is easy to see that g is a bounded continuous extension of f to the domain S. According to the definition of $\beta(S)$, there exists a continuous extension φ of g (hence of f) to the domain $\beta(S)$. We have

$\varphi(x) = f(x) = 0$ for each $x \epsilon \Phi_1$ and $\varphi(x) = f(x) = 1$ for each $x \epsilon \Phi_2$. Since φ is continuous, we must have $\varphi(x) = 0$ for each $x \epsilon \bar{\Phi}_1$ and $\varphi(x) = 1$ for each $x \epsilon \bar{\Phi}_2$, so that $\bar{\Phi}_1\bar{\Phi}_2 = 0$, which is a contradiction.

The topological space S will be called *locally normal* if each point $x \epsilon S$ possesses a neighborhood U such that $\bar{U}$ is a normal space. Any normal space is locally normal; more generally, any open subset of a locally normal space is locally normal.

A locally normal Riesz space S is completely regular. Let a be a given point of a locally normal space S and let V be a given neighborhood of a. There exists a neighborhood U of a such that $\bar{U}$ is a normal space. Also $\overline{UV}$ is a normal space, since it is a closed subset of $\bar{U}$. Since (a) and $\overline{UV} - UV$ are two closed subsets of the normal space $\overline{UV}$ without a common point, there exists a continuous function f in the domain $\overline{UV}$ such that $f(a) = 0$ and $f(x) = 1$ for each $x \epsilon \overline{UV} - UV$. For $x \epsilon S$ put (i) $g(x) = f(x)$ if $x \epsilon UV$, (ii) $g(x) = 1$ if $x \epsilon S - UV$. Then it is easy to see that g is a continuous function in the domain S such that $g(a) = 0$ and $g(x) = 1$ for each $x \epsilon S - V$. Therefore S is completely regular.

A completely regular space S need not be locally normal. Let ω be the first infinite ordinal number. Let ω_1 be the first uncountable ordinal number. Let S_1 be the space of all ordinal numbers $\leqq \omega$. Let S_2 be the space of all ordinal numbers $\leqq \omega_1 \cdot \omega$. The topology in S_1 and in S_2 is defined in the usual way by means of intervals. Let S_{12} be the cartesian product of the two spaces S_1 and S_2. Let T be the set of all points $(x, y) \epsilon S_{12}$, for which $x = \omega$ and $y = \omega_1 \cdot n (n = 1, 2, 3, \cdots)$. Let $S = S_{12} - T$. Then S is a completely regular space, but it is not locally normal.

It is easy to see that a completely regular space S is locally normal if and only if $S \subset \lambda(S)$. I do not know whether there exists a completely regular space $S \neq 0$ such that $S \cdot \lambda(S) = 0$.

A Riesz space S is locally normal if and only if it is homeomorphic with an open subset of a normal Riesz space.[19] We know that an open subset of a normal Riesz space is a locally normal Riesz space. Conversely let S be a locally normal Riesz space. Let S_0 be a new space consisting of all points of S and of a single new point ω. The topology of S_0 is defined as follows. If $\omega \epsilon A \subset S_0$, then A is closed in S_0 if and only if $A - (\omega)$ is closed in S. If $A \subset S_0 - (\omega) = S$, then A is closed in S_0 if and only if (i) A is closed in S, (ii) $\bar{A} \subset \lambda(S)$, the bar indicating closure in $\beta(S)$. It is easy to see that S_0 is a Riesz space and that S is an open subset of S_0. It remains to be shown that the space S_0 is normal. Let F_1 and F_2 be two closed subsets of S_0 such that $F_1F_2 = 0$. Since the point ω belongs at most to one of the two sets F_1 and F_2, we may suppose that $F_1 \subset S$. Since F_1 is closed in S_0, the closure $\bar{F}_1$ of F_1 in the space $\beta(S)$ is a subset of $\lambda(S)$. Put $F_3 = F_2 - (\omega)$. Then F_1 and F_3 are two closed subsets of S and $F_1F_3 = 0$. We know that $\bar{F}_1 \cdot \bar{F}_3 \cdot \lambda(S) = 0$ (the closures being formed again in $\beta(S)$). But

[19] I do not know whether the restriction to Riesz spaces is really necessary in this theorem.

$\bar{F}_1 \subset \lambda(S)$ so that $\bar{F}_1$ and $\bar{F}_3 + \beta(S) - \lambda(S)$ are two closed subsets of $\beta(S)$ without a common point. Since $\beta(S)$ is a bicompact Hausdorff space, it is normal, so that there exists a continuous function φ in the domain $\beta(S)$ such that $\varphi(x) = 0$ for each $x \in \bar{F}_1$ and $\varphi(x) = 1$ for each $x \in \bar{F}_3$ and for each $x \in \beta(S) - \lambda(S)$. Let us define a function f in the domain S_0 in the following way. If $x \in S$, then $f(x) = \varphi(x)$; moreover $f(\omega) = 1$. Then it is easy to see that f is a continuous function in the domain S_0 such that $f(x) = 0$ for each $x \in F_1$ and $f(x) = 1$ for each $x \in F_2$.

I conclude with two more unsolved questions. A topological space S is called *completely normal* if every subset of S is a normal space. S may be called *locally completely normal* if every point $x \in S$ possesses a neighborhood U such that $\bar{U}$ is a completely normal space. S may be called *completely locally normal* if every subset of S is a locally normal space. It is easy to see that a locally completely normal space is completely locally normal. I do not know whether the converse holds true. Any open subset of a completely normal space is a locally completely normal space. I do not know whether a locally completely normal space must be homeomorphic with an open subset of a completely normal space.

BRNO, CZECHOSLOVAKIA.

Annals of Mathematics
Vol. 38, No. 4, October, 1937

REMARK ON BICOMPACT SPACES

By Bedřich Pospíšil

(Received April 26, 1937)

We write $\exp \mathfrak{x} = 2^{\mathfrak{x}}$ for each cardinal number $\mathfrak{x}$.

Theorem I. *For each infinite cardinal number* $\mathfrak{h}$, *there exists a bicompact Hausdorff space* S *and a subspace* $T \subset S$ *such that* (i) T *is an isolated space,* (ii) T *is dense in* S, (iii) *the cardinal number of* T *is* $\mathfrak{h}$, (iv) *the cardinal number of* S *is* $\exp \exp \mathfrak{h}$.

Proof. Let H denote a set of $\mathfrak{h}$ elements. Let X denote the set of all functions φ defined over H and assuming only values 0 and 1. Let Ω denote the family of all subsets of X such that any $A \in \Omega$ is the set of all functions $\varphi \in X$ assuming a given value at each of a given finite number of given elements of H. Then the cardinal number of X is $\exp \mathfrak{h}$ and the cardinal number of Ω is $\mathfrak{h}$. We consider X as a topological space, a neighborhood of $\varphi \in X$ being any $A \in \Omega$ such that $\varphi \in A$. It is easy to see that we may choose a set E dense in X of cardinal number $\mathfrak{h}$. Let F denote the set of all mappings of X into a subset of X. Let T denote the set of all finite families t of pairs (C_k, e_k) where $C_k \in \Omega$, $e_k \in E$ and any two sets C_k have a vacuous intersection; the pairs (C_k, e_k) will be termed the *coordinates* of $t \in T$.

We put $S = T + F$, the topology of S being defined by neighborhoods as follows. A neighborhood of $t \in T$ consists of t only. Neighborhoods of $f \in F$ will be defined in a somewhat complicated manner. Choose a finite number of different points $x_k \in X (1 \leqq k \leqq n)$. For $1 \leqq k \leqq n$, choose sets $A_k \in \Omega$ and $B_k \in \Omega$ such that $x_k \in A_k$ and $f(x_k) \in B_k$. Let Q be the set of all $g \in F$ such that $g(x_k) \in B_k$ for $1 \leqq k \leqq n$. Define a set $P \subset T$ in the following manner. An element t of T belongs to P if and only if, for $1 \leqq k \leqq n$, there exists a coordinate (C_k, e_k) of t such that $x_k \in C_k$, $C_k \subset A_k$, $e_k \in B_k$. [Of course, t may possess *more* than n coordinates.] Let K be any finite subset of P. Put $O = P - K + Q$. Any such set O is a neighborhood of f in our space S.

The properties (i)-(iv) being evident, to prove theorem I it suffices to show that S is a bicompact Hausdorff space. This is indeed true but a quicker way of completing our proof is as follows. It is easy to see that all our neighborhoods are both closed and open in our space S, whence it readily follows that S is a completely regular space. Therefore, using a result of Tychonoff,[1] there exists a bicompact Hausdorff space B containing S. If S_0 denotes the closure of S in B, then the spaces S_0 and $T \subset S_0$ satisfy our theorem. Indeed, it is easy to prove

[1] A. Tychonoff, *Über die topologische Erweiterung von Räumen*, Math. Annalen, 102, 1930, 544–561.

that the cardinal number of S_0 cannot exceed exp exp $\mathfrak{h}$, since S_0 is a Hausdorff space containing a dense subset of cardinal number $\mathfrak{h}$.

In a recent paper,[2] Čech attached to any completely regular space T a unique bicompact space $\beta(T) \supset T$ and posed the problem of determining the cardinal number of $\beta(T)$ in case T is an isolated countable space. The answer to Čech's question is a particular case of the following

THEOREM II. *Let T denote an infinite isolated space of cardinal number $\mathfrak{h}$; then the cardinal number of $\beta(T)$ is* exp exp $\mathfrak{h}$.

PROOF. Again applying the final sentence in the proof of theorem I, the cardinal number of $\beta(T)$ cannot exceed exp exp $\mathfrak{h}$. By theorem I, there exists a bicompact space $S \supset T$ of cardinal number exp exp $\mathfrak{h}$ such that T is dense in S. In his paper, already quoted, Čech proved that there exists a continuous mapping of $\beta(T)$ into S. Hence the cardinal number of $\beta(T)$ is not less than exp exp $\mathfrak{h}$.

BRNO, CZECHOSLOVAKIA.

[2] E. Čech, *On bicompact spaces*, Annals of Math., 38, 1937, 823–844.

Annals of Mathematics
Vol. 38, No. 4, October, 1937

ZUM HOPFSCHEN UMKEHRHOMOMORPHISMUS

Von Hans Freudenthal

(Received January 11, 1937)

Eine Abbildung einer Mannigfaltigkeit auf eine andere induziert einen Homomorphismus zwar der Bettischen Gruppen der einen in die der andern, nicht aber im Allgemeinen auch des einen Schnittringes in den andern. Vor längerer Zeit hat aber H. Hopf [1] mit der Lefschetzschen Produktmethode "Umkehrungen" von Mannigfaltigkeitsabbildungen angegeben; eine solche Umkehrung besitzt unter anderen wichtigen Eigenschaften die, ein Schnittringhomomorphismus zu sein; aus diesen Eigenschaften hat H. Hopf eine Reihe bedeutsamer Folgerungen gezogen. Später hat H. Hopf [2] zur Untersuchung der Abbildungen der S_3 auf die S_2 wiederum derartige Umkehrungen definiert (für Abbildungen verschieden-dimensionaler Mannigfaltigkeiten). Auch Verf. [4] hat sich mit derartigen Umkehrungen beschäftigt, und zwar im Außenraum abgeschlossener Teilmengen von Sphären.

Hier soll gezeigt werden,[1] wie sich diese Umkehrhomomorphismen sehr einfach im Rahmen der neueren Homologietheorie ergeben; übrigens wird sich ihr Gültigkeitsbereich einigermaßen erweitern. Alle diese "Umkehrungen" kommen übrigens auf dasselbe hinaus; die Art, wie sie hier eingeführt werden, erinnert wohl sehr an die Methode von H. Hopf [2] und Verf. [4], ist aber vielleicht noch etwas einfacher. Die Haupteigenschaft der "Umkehrungen" ergibt sich aber sicher viel einfacher als bisher.

Eine frühere Arbeit des Verfassers[1a] wird als bekannt vorausgesetzt und mit AG zitiert.

1. Wir übernehmen die Bezeichungen von AG. Insbesondere sei also $\mathfrak{d}$ der Dualitätsoperator, der (in d-dimensionalen Homomlogiemannigfaltigkeiten) B^p topologisch isomorph auf B_{d-p} abbildet. Wir nehmen immer eine bestimmte Eckenordnung an und können dann praktischer mit $\mathfrak{o}\mathfrak{d}$ arbeiten. Wir hatten (AG(2)):

$$\mathfrak{o}\mathfrak{d}[a_0 \cdots a_p] = \sum \pi(a_0 \cdots a_p \cdots a_d)[a_p \cdots a_d]$$

(in allen derartigen Formeln seien nur *normale* Simplexe zugelassen).

[1] Zusatz bei der Korrektur: Unsere Ergebnisse überschneiden sich mit denen von H. Whitney in einer inzwischen erschienenen Note: On products in a complex, Proc. Nat. Acad. USA **23** (1937), 285–291.

[1a] *Alexanderscher und Gordonscher Ring und ihre Isomorphie*, Ann. of Math. **36** (1937), 647–655.

Da für die Bettischen Gruppen $\mathfrak{d}$ ein Isomorphismus ist, so ist auch das durch

$$(z^{d-p}\mathfrak{d})z^p = z^{d-p}(\mathfrak{d}z^p)\ ^2$$

definierte hintere $\mathfrak{d}$ ein Isomorphismus und zwar, wie sich zeigen wird, im Wesentlichen derselbe. Dagegen wird natürlich keineswegs auch für die Gruppe der Komplexe vorderes und hinteres $\mathfrak{o}\mathfrak{d}$ identisch sein. Man hat vielmehr

$$[b_\rho \cdots b_d](\mathfrak{o}\mathfrak{d}[a_0 \cdots a_\rho]) = \sum \pi(a_0 \cdots a_\rho b_{\rho+1} \cdots b_d),$$

falls diese Ecken ein Simplex erzeugen und $a_\rho = b_\rho$ ist, sonst 0; die rechte Seite ist aber auch gleich

$$\sum \pi(b_0 \cdots b_d)[b_0 \cdots b_\rho][a_0 \cdots a_\rho]$$

(bei festem $[b_\rho \cdots b_d]$ zu summieren) bzw. 0. Also

$$[b_\rho \cdots b_d]\ \mathfrak{o}\mathfrak{d} = \sum \pi(b_0 \cdots b_d)[b_0 \cdots b_\rho].$$

Hinteres $\mathfrak{o}\mathfrak{d}$ liefert also für die Komplexe dasselbe wie vorderes $\mathfrak{o}'\mathfrak{d}$ (wenn $\mathfrak{o}'$ die simpliziale Verschiebung ist, die auf der umgekehrten Anordnung der Ecken beruht); für die Bettischen Gruppen unterscheiden vorderes und hinteres $\mathfrak{d}$ sich also in der Tat nicht:

(D) $$\mathfrak{d}z^p \sim z^p\mathfrak{d}$$

2. Wir hatten weiter das *Alexandersche Produkt* (für normale obere Simplexe) $[a_0 \cdots a_\rho]\cdot[a_\rho \cdots a_{\rho+\sigma}] = [a_0 \cdots a_\rho \cdots a_{\rho+\sigma}]$, sonst 0, ferner in d-dimensionalen Homologiemannigfaltigkeiten den *Schnitt* (AG(4)) $\mathfrak{d}t^\rho \times t_\sigma$ oder praktischer $\mathfrak{o}(\mathfrak{d}t^\rho \times t_\sigma)$:

$$\mathfrak{o}(\mathfrak{d}[a_0 \cdots a_\rho] \times [a_0 \cdots a_\sigma]) = [a_\rho \cdots a_\sigma];^3$$

schließlich im Außenraum $S\backslash A$ eines Teilpolytops A der d-dimensionalen Homologiesphäre das *Gordonsche Produkt*

$$z_\rho \otimes z_\sigma = \mathfrak{r}(k_{\rho+1} \times k_{\sigma+1}) \text{ mit } z_\rho = \mathfrak{r}k_{\rho+1},\ z_\sigma = \mathfrak{r}k_{\sigma+1} \text{ (in } S).$$

Wir hatten den Isomorphismus $\mathfrak{a} = \mathfrak{r}\mathfrak{d}$ oder, praktischer,

$$\mathfrak{a} = \mathfrak{r}\mathfrak{o}\mathfrak{d}$$

von $B^\rho(A)$ auf $B_{d-\rho-1}(S\backslash A)$ und (AG (5)) für $z^\rho, z^\sigma \subset A$:

$$\mathfrak{a}(z^\rho \cdot z^\sigma) \sim \mathfrak{a}z^\sigma \otimes \mathfrak{a}z^\rho \text{ in } S\backslash A.$$

---

[2] Die hier gebrauchte Produktbildung zwischen oberen und unteren Komplexen usw. gleicher Dimension (siehe AG 1) setzt immer duale Koeffizientenbereiche voraus. Man verwechsle sie nicht mit der Alexanderschen (AG 10), die immer mit einem Multiplikationspunkt geschrieben wird (der *hier* fehlt).

[3] Wir haben leider in AG bei der Schnittdefinition beide Faktoren vertauscht, was nur einen Vorzeichenunterschied gegen die übliche Definition ergibt, aber sonst kaum etwas ausmacht. Trotz der kleinen Unsymmetrien, die entstehen, behalten wir die Definition bei. Die Formeln (U'), (A'), (V') usw. werden nach der Vertauschungsformel für den Schnitt von dieser Abweirhung übrigens nicht beeinflußt.

Endlich hat man nach E. Čech [1] (in beliebigen d-dimensionalen Homologiemannigfaltigkeiten) die Isomorphie

$$\mathfrak{d}(z^\rho \cdot z^\sigma) \sim \mathfrak{d}z^\sigma \times \mathfrak{d}z^\rho,$$

die sich bei uns so ergibt: Es genügt,

$$\mathfrak{o}\mathfrak{d}(z^\rho \cdot z^\sigma) \sim \mathfrak{o}(\mathfrak{d}z^\sigma \times \mathfrak{o}\mathfrak{d}z^\rho)$$

zu beweisen. Das ergibt sich wiederum aus

$$\mathfrak{o}\mathfrak{d}(t^\rho \cdot t^\sigma) = \mathfrak{o}(\mathfrak{d}t^\sigma \times \mathfrak{o}\mathfrak{d}t^\rho),$$

was wir folgendermaßen verifizieren: Sei $t^\rho = [a_0 \cdots a_\rho]$; dann ist die linke Seite gleich $\sum \pi(a_0 \cdots a_\rho \cdots a_{\rho+\sigma} \cdots a_d)[a_{\rho+\sigma} \cdots a_d]$, falls $t^\sigma = [a_\rho \cdots a_{\rho+\sigma}]$ ist, sonst 0. Rechts hat man für solche Wahl von t^ρ und t^σ: $\mathfrak{o}\mathfrak{d}t^\rho = \sum \pi(a_0 \cdots a_\rho \cdots a_d)[a_\rho \cdots a_d]$. Beide Seiten stimmen also überein.

Damit hat sich die Isomorphie beider Produktbildungen ergeben.

3. Wir betrachten jetzt eine stetige Abbildung $\mathfrak{f}$ einer μ-dimensionalen Homologiemannigfaltigkeit M in eine ν-dimensionale, N. Ohne Beschränkung der Allgemeinheit dürfen wir voraussetzen, daß $\mathfrak{f}$ simplizial ist und die Reihenfolge der Ecken nicht stört (also $\mathfrak{o}\mathfrak{f} = \mathfrak{f}\mathfrak{o}$).

Bei den Operationen $\mathfrak{o}$ und $\mathfrak{d}$ müssen wir im Folgenden unterscheiden, ob sie in M oder in N genommen sind; da im Allgemeinen keine Mißverständnisse möglich sind, wollen wir diesen Unterschied nicht explizit ausdrücken; nur wo es unbedingt nötig ist, hängen wir den Index M oder N an.

Gemäß AG 2 induziert $\mathfrak{f}$ Homomorphismen

$$\mathfrak{f}K_\rho(M) \subset K_\rho(N),\ \mathfrak{f}B_\rho(M) \subset B_\rho(N),$$

$$K^\rho(N)\mathfrak{f} \subset K^\rho(M),\ B^\rho(N)\mathfrak{f} \subset B^\rho(M).$$

Bei Homologiemannigfaltigkeiten ist aber auch

$$\mathfrak{f}^* = \mathfrak{d}^{-1}\mathfrak{f}\mathfrak{d}$$

ein Homomorphismus

$$\mathfrak{f}^* B^\rho(M) \subset B^{\rho-\mu+\nu}(N),$$

$$B_\rho(N)\mathfrak{f}^* \subset B_{\rho+\mu-\nu}(M).$$

$\mathfrak{f}^*$ wirkt also gerade in umgekehrter Richtung wie $\mathfrak{f}$.

Man hat die Relationen

$$\text{(R)} \qquad \mathfrak{d}\mathfrak{f}^* = \mathfrak{f}\mathfrak{d}, \quad \text{oder praktischer} \qquad \mathfrak{o}\mathfrak{d}\mathfrak{f}^* = \mathfrak{f}\mathfrak{o}\mathfrak{d},$$

was auch für Komplexe sinnvoll ist.

Es ist für die Anwendungen sehr wichtig, daß $\mathfrak{f}^*$ nicht einfach abstrakt definiert ist, sondern daß es etwa zu einem unteren Zyklus in N ganz konkret ein $\mathfrak{f}^*$-Bild liefert. In Worten ausgedrückt lautet die Konstruktion z. B. so:

Man nehme den Zyklus als Linearform von Dualzellen an; von einer Dualzelle erhält man das $\mathfrak{f}^*$ -Bild, indem man vom zugehörigen Simplex der ursprünglichen Teilung die (gleichdimensionalen) Urbilder sucht und deren Dualzellen addiert.

Das hintere $\mathfrak{f}^*$ ist damit für untere Komplexe aus Dualzellen ($\mathfrak{d}k^{\nu-\rho}$) erklärt und bildet $\mathfrak{d}K^{\nu-\rho}(N)$ homomorph in $\mathfrak{d}K^{\nu-\rho}(M)$ ab. Nach AG (1) geht dabei der Rand eines Komplexes in den Rand des Bildes über, also ein Zyklus in einen Zyklus, usw., so daß diese konkrete Definition (unter Berücksichtigung von (D)) dasselbe liefert wie die obige abstrakte (für die Elemente der Bettischen Gruppen).

Man bemerkt, daß die zu $\mathfrak{f}\mathfrak{f}^*z_\rho$ gehörige Punktmenge in der zu z_ρ gehörigen Punktmenge enthalten ist oder wenigstens in denselben ν-dimensionalen Simplexen wie die von z_ρ liegt. Das ist für Anwendungen (H. Hopf [2]) sehr wichtig und würde bereits die Auffassung von $\mathfrak{f}^*$ als "Umkehrung" von $\mathfrak{f}$ rechtfertigen.

Es gilt aber noch mehr: Wir werden beweisen:

(U) $$\mathfrak{f}^*(z^\rho\mathfrak{f}) \sim \gamma z^\rho \text{ für } \mu \geqq \nu,$$

(γ = Abbildungsgrad von $\mathfrak{f}$).

Weiter wissen wir (AG, 10), daß das hintere $\mathfrak{f}$ ein Ringhomomorphismus ist,

(A) $$(z^\rho \cdot z^\sigma)\mathfrak{f} \sim z^\rho\mathfrak{f}\cdot z^\sigma\mathfrak{f};$$

etwas Derartiges können wir von seiner "Umkehrung", dem vorderen $\mathfrak{f}^*$, natürlich nicht erwarten, wohl aber erhält man, wenn man in (U) z^ρ durch $z^\rho\cdot z^\sigma$ ersetzt und (A) berücksichtigt,

$$\mathfrak{f}^*(z^\rho\mathfrak{f}\cdot z^\sigma\mathfrak{f}) \sim z^\rho\cdot\gamma z^\sigma \sim z^\rho\cdot\mathfrak{f}^*(z^\sigma\mathfrak{f}).$$

Wir werden aber viel mehr beweisen: ganz allgemein (für beliebige μ, ν) gilt

(V) $$\mathfrak{f}^*(z^\rho\cdot z^\sigma\mathfrak{f}) \sim \mathfrak{f}^*z^\rho\cdot z^\sigma$$

(hier ist natürlich $z^\rho \subset M$, $z^\sigma \subset N$).

(U) ist (wenigstens für $\mu = \nu$) ein Spezialfall von (V), den man erhält, wenn man in (V) für z^ρ den 0-dimensionalen oberen Grundzyklus von M (die Summe der Ecken), also für $\mathfrak{f}^*z^\rho$ den γ-fachen 0-dimensionalen Grundzyklus von N einsetzt.

Nehmen wir vorläufig (U) und (V) als bewiesen an! Ersetzen wir in (U), (A), (V) bzw. z^ρ und z^σ durch $z_{\rho'}\mathfrak{d}$ und $z_{\sigma'}\mathfrak{d}$ (was wir bis auf Homologien dürfen), berücksichtigen wir (R), (D) und die Beziehung, die nach 2 zwischen Produkt- und Schnittbildung besteht, und schreiben wir zum Schluß wieder ρ und σ für ρ' und σ', so erhalten wir

(U′) $$\mathfrak{f}(z_\rho\mathfrak{f}^*) \sim \gamma z_\rho \quad \text{für } \mu \geqq \nu,$$

(A′) $$(z_\rho \times z_\sigma)\mathfrak{f}^* \sim z_\rho\mathfrak{f}^* \times z_\sigma\mathfrak{f}^*,$$

(V′) $$\mathfrak{f}(z_\rho\mathfrak{f}^* \times z_\sigma) \sim z_\rho \times \mathfrak{f}z_\sigma.$$

Umgekehrt kann man natürlich aus (U′), (A′), (V′) wieder (U), (A), (V) ableiten.

(U′), (A′), (V′) sind dieselben Formeln, die bei der Hopfschen Umkehrung [1] auftreten. Man darf daher vermuten, daß das hintere $\mathfrak{f}^*$ selbst mit der Hopfschen Umkehrung übereinstimmt; in 7 werden wir das auch beweisen. Dann haben wir für die Hopfsche Umkehrung eine neue (einfachere) Definition gegeben und ihre Haupteigenschaften von neuem (einfacher) abgeleitet.

Seien nun die Mannigfaltigkeiten M und N Homologiesphären R und S, und sei $\mathfrak{f}R \subset S$, $\mathfrak{f}A \subset B$ (A und B abgeschlossene Teilmengen von R und S). Eine einfache Überlegung wird aus (U′), (A′), (V′) die Formeln (U″), (A″), (V″) ergeben:

(U″) $\qquad \mathfrak{f}(z_\rho \mathfrak{f}^*) \sim \gamma z_\rho \quad \text{in } S\backslash B \text{ für } \mu \geqq \nu,$

(A″) $\qquad (z_\rho \otimes z_\sigma)\mathfrak{f}^* \sim z_\rho \mathfrak{f}^* \otimes z_\sigma \mathfrak{f}^* \quad \text{in } R\backslash A,$

(V″) $\qquad \mathfrak{f}(z_\rho \mathfrak{f}^* \otimes z_\sigma) \sim z_\rho \otimes \mathfrak{f} z_\sigma \quad \text{in } S\backslash B.$

4. Wir beweisen nun (U′) und (V′) (statt (U) und (V)). Wir nehmen in (U′) $z^{\nu-\rho}\mathfrak{o}\mathfrak{d}$ statt z_ρ und haben dann (wegen (R))

$$\mathfrak{f}(z^{\nu-\rho}\mathfrak{o}\mathfrak{f}\mathfrak{d}) \sim \gamma z^{\nu-\rho}\mathfrak{o}\mathfrak{d} \text{ für } \mu \geqq \nu$$

zu beweisen. Wir werden sogar

$$\mathfrak{f}(z^{\nu-\rho}\mathfrak{f}\mathfrak{o}\mathfrak{d}) = \gamma z^{\nu-\rho}\mathfrak{o}\mathfrak{d}$$

beweisen,[4] und das ergibt sich aus

$$\mathfrak{f}(t^{\nu-\rho}\mathfrak{f}\mathfrak{o}\mathfrak{d}) = \gamma t^{\nu-\rho}\mathfrak{o}\mathfrak{d},$$

was wir jetzt verifizieren: Sei $t^{\nu-\rho} = [b_\rho \cdots b_\nu]$. Dann ist $t^{\nu-\rho}\mathfrak{f} = \sum [a_\rho \cdots a_\nu]$ mit $b_\rho = \mathfrak{f}a_\rho, \cdots, b_\nu = \mathfrak{f}a_\nu$, also $t^{\nu-\rho}\mathfrak{f}\mathfrak{o}\mathfrak{d} = \sum \pi(a_0 \cdots a_\rho \cdots a_\nu)\,[a_0 \cdots a_\rho]$ erstreckt über alle $[a_0 \cdots a_\rho \cdots a_\nu]$ mit $b_\rho = \mathfrak{f}a_\rho, \cdots, b_\nu = \mathfrak{f}a_\nu$. Links haben wir demnach

$$\mathfrak{f}(t^{\nu-\rho}\mathfrak{f}\mathfrak{o}\mathfrak{d}) = \sum \pi(a_0 \cdots a_\rho \cdots a_\nu)[b_0 \cdots b_\rho] = \sum [b_0 \cdots b_\rho] \sum \pi(a_0 \cdots a_\rho \cdots a_\nu),$$

wo $b_0 = \mathfrak{f}a_0, \cdots, b_\rho = \mathfrak{f}a_\rho$ gesetzt ist und die innere Summe sich auf festes $[b_0 \cdots b_\rho]$ bezieht, während die äußere über alle $[b_0 \cdots b_\rho]$ zu erstrecken ist, die mit $[b_\rho \cdots b_\nu]$ ein ν-dimensionales Simplex bilden. Die innere Summe ist aber gleich $\gamma\pi(b_0 \cdots b_\rho \cdots b_\nu)$, also stimmen linke und rechte Seite tatsächlich überein.

Führen wir in (V′) die Substitution $z_\rho = z^{\nu-\rho}\mathfrak{d}$ aus, so haben wir nur (unter Berücksichtigung von (D))

$$\mathfrak{f}\mathfrak{o}(\mathfrak{d}(t^{\nu-\rho}\mathfrak{f}) \times t_\sigma) = \mathfrak{o}(\mathfrak{d}t^{\nu-\rho} \times \mathfrak{f}t_\sigma)$$

zu beweisen. Nehmen wir $t_\sigma = [a_0 \cdots a_\sigma]$, $\mathfrak{f}t_\sigma = [b_0 \cdots b_\sigma]$, so verschwindet gemäß 2 die rechte Seite dann und nur dann nicht, wenn $t^{\nu-\rho} = [b_0 \cdots b_{\nu-\rho}]$ ist; sie wird dann $[b_{\nu-\rho} \cdots b_\sigma]$. Links erhält man dann und nur dann etwas Nicht-

[4] Man beachte $\mathfrak{o}\mathfrak{f} = \mathfrak{f}\mathfrak{o}$.

verschwindendes—und zwar auch wieder $[b_{\nu-\rho} \cdots b_\sigma]$—, wenn in $t^{\nu-\rho}\mathfrak{f}$ $[a_0 \cdots a_{\nu-\rho}]$ auftritt, also auch wieder für $t^{\nu-\rho} = [b_0 \cdots b_{\nu-\rho}]$.

Um (U''), (A''), (V'') zu erhalten, müssen wir etwas genauer verfahren. Wir brauchen uns nur um Polytope A und B zu kümmern. Die Ecken von S dürfen wir so anordnen, daß die von B hinter allen andern kommen. Ist nun z_ρ ein Zyklus aus $S\backslash B$, so kann man $z^{\nu-\rho}\mathfrak{o}\mathfrak{d}$ innerhalb $S\backslash B$ wählen, so daß es in $S\backslash B$ mit z_ρ homolog ist. Analog darf man mit einem etwaigen $k_{\rho+1}$ aus $S\backslash B$ verfahren, dessen Rand $z_{\rho+1}$ ist. Wendet man darauf hinten $\mathfrak{f}^*$ an, so kommt man in die Urbildmenge von $S\backslash B$, also sicherlich in $R\backslash A$, woran die weitere Anwendung von $\mathfrak{o}\mathfrak{d}$ nichts ändert. $\mathfrak{f}^*$ ist damit als Homomorphismus

$$B_\rho(S\backslash B)\mathfrak{f}^* \subset B_{\rho+\mu-\nu}(R\backslash A)$$

definiert. Die Gleichheit, die wir oben im Beweise von (U') erhalten haben, zeigt aber unmittelbar, daß die durch (U'') ausgedrückte Homologie wirklich in $S\backslash B$ gilt. (A'') ist ohne weiteres klar, und (V'') ergibt sich analog aus (U'') bei genauerer Betrachtung des Beweises von (U').

5. Die folgenden Sätze beziehen sich wieder auf beliebige Homologiemannigfaltigkeiten.

a. $(\mathfrak{f}\mathfrak{g}) = \mathfrak{f}^*\mathfrak{g}^*$. (Beweis klar.)

b. Ist die Homologiemannigfaltigkeit M Teilmenge von N und $\mathfrak{g}$ die Abbildung, die M punktweise auf sich selbst abbildet, so ist $z_\rho\mathfrak{g}^* \sim z_\rho \times M$ (wo M auch den μ-dimensionalen Grundzyklus von M bezeichne).—Denn wenn wir in (U') z_σ durch M, also $\mathfrak{g}z_\sigma$ auch durch M ersetzen, können wir links den Faktor M weglassen und erhalten die gewünschte Beziehung.

c. Ist $\mathfrak{f}M \subset N$, $\mathfrak{h}M' \subset N$, ist weiter M' Teilmannigfaltigkeit von M und stimmen $\mathfrak{f}$ und $\mathfrak{h}$ auf M' überein, so ist $z_\rho\mathfrak{h}^* \sim z_\rho\mathfrak{f}^* \times M'$.—Das folgt aus a und b, wenn man noch die Abbildung $\mathfrak{g}$ $M' \subset M$ einführt, die auf M' die Identität ist.

d. Ist $M \subset N$, so ist es gleichgültig, ob man den Schnitt zweier Zyklen aus M in M oder in N bildet.—Folgt aus (A').

e. (;) bedeute die Bildung des Cartesischen Produktes. $\mathfrak{p}$ sei die Projektion von $(M;N)$ auf N. Dann ist (für $z_\rho \subset N$):

$$(M; z_\rho) \sim z_\rho\mathfrak{p}^*.$$

Beim Beweise machen wir von einer anderwärts von Verf. [5] abgeleiteten Simplizialzerlegung des Cartesischen Produktes zweier Simplexe t_μ und t_ν Gebrauch. Die Ecken von t_μ werden mit $0, \cdots, \mu$, die von t_ν mit $0, \cdots, \nu$ numeriert. Ecken von $(t_\mu; t_\nu)$ sind die Paare ($\alpha\beta$ aus den Ecken von t_μ und t_ν; man kann sie sich in der Matrix

$$\mathfrak{M}: \quad \begin{matrix} 00 & 01 & \cdots & 0\nu \\ 10 & 11 & \cdots & 1\nu \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \mu 0 & \mu 1 & \cdots & \mu\nu \end{matrix}$$

aufschreiben Die $(\mu + \nu)$-dimensionalen Simplexe $u_{\mu+\nu}$ von $(t_\mu; t_\nu)$ werden so definiert: Jedes $u_{\mu+\nu}$ besteht aus einer Folge von Elementen von $\mathfrak{M}$, die mit 00 beginnt, mit $\mu\nu$ endet, und in der jedes Element rechter oder unterer Nachbar seines Vorgängers ist. $1 + \psi_\beta(u_{\mu+\nu})$ sei die Zahl der Elemente von $u_{\mu+\nu}$ in der β-ten Spalte; $\eta(u_{\mu+\nu}) = \sum \beta\psi_\beta(u_{\mu+\nu})$. Addiert man diese $u_{\mu+\nu}$ mit dem Koeffizienten $(-1)^{\eta(u_{\mu+\nu})}$, so erhält man die gewünschte Simplizialzerlegung von $(t_\mu; t_\nu)$.

Um

$$(M; z_\rho) \sim z_\rho \mathfrak{p}^*$$

zu beweisen, nehmen wir uns je ein willkürliches Simplex t_μ aus M und t_ν aus N und berechnen auf beiden Seiten den in $(t_\mu; t_\nu)$ liegenden Anteil. Wir können das auch so ausdrücken: Wir ersetzen M durch t_μ und N durch t_ν und beweisen

$$(t_\mu; t_\rho) = t^{\nu-\rho}\mathfrak{p}\mathfrak{o}\mathfrak{d} \text{ mit } t_\rho = t^{\nu-\rho}\mathfrak{o}\mathfrak{d}_N .$$

Sind die Ecken von t_μ und t_ν wie oben numeriert, so verschwindet die linke Seite dann nur dann nicht, wenn $t_\rho = [0 \cdots \rho]$, also $t^{\nu-\rho} = [\rho \cdots \nu]$ ist; sie wird dann gleich der algebraischen Simplexsumme, gebildet aus der Matrix $\mathfrak{M}'$, die nur die 0-te bis ρ-te Spalte von $\mathfrak{M}$ umfaßt. Rechts hat man $t^{\nu-\rho}\mathfrak{p} = \sum [\alpha_\rho\rho \cdots \alpha_\nu\nu]$; zu $t^{\nu-\rho}\mathfrak{p}\mathfrak{o}\mathfrak{d}$ liefert aber nur $[\mu\rho \cdots \mu\nu]$ einen Beitrag, und zwar gerade das, was wir links ausgerechnet haben. Für andere Wahlen von t_ρ verschwindet auch die rechte Seite, so daß unsere Gleichung bewiesen ist.

6. Wir zeigen nun, daß die Hopfsche Umkehrung [1] von Mannigfaltigkeitsabbildungen mit unserer übereinstimmt.

Die Hopfsche Umkehrung war so erklärt: Man bilde in $(M; N)$ das Lefschetzsche "Bild" von $\mathfrak{f}$, d.h. die Mannigfaltigkeit M', die aus den Punkten $(a; \mathfrak{f}a)$ zusammengesetzt ist. $(M; z_\rho) \times M'$ ist ein Zyklus in M' (wenn z_ρ einer in N ist); seine Projektion auf M ist das Umkehrungsbild von z_ρ. Wenn man M' als einen Repräsentanten von M auffaßt und $\mathfrak{f}$ durch die Projektion von M' auf N ersetzt, kann man sich die letzte Projektion übrigens sparen.

Nun ist nach 5e: $(M; z_\rho) \sim z_\rho \mathfrak{p}^*$, also $(M; z_\rho) \times M' \sim z_\rho\mathfrak{p}^* \times M'$. Das ist aber nach 5c und der letzten Bemerkung des vorigen Absatzes wesentlich nichts Anderes als $z_\rho\mathfrak{f}^*$. Damit ist die Äquivalenz bewiesen.

LITERATUR

siehe AG, außerdem:

H. Freudenthal: 4. *Über die topologische Invarianz kombinatorischer Eigenschaften des Außenraumes abgeschlossener Mengen.* Compositio Math. *2* (1935), 163–176.

5. *Eine Simplizialzerlegung des Cartesichen Produktes zweier Simplexe.* Fundamenta Math. *29* (1937), 138–144.

H. Hopf: 1. *Zur Algebra der Abbildungen von Mannigfaltigkeiten.* Journ. f. d. r. u. angew. Math. *163* (1930), 71–88.

2. *Über die Abbildungen der dreidimensionalen Sphäre auf die Kugelfläche.* Math. Ann. *104* (1931), 637–665.

Amsterdam, Holland.

Annals of Mathematics
Vol. 38, No. 4, October, 1937

NOTE ON ALGEBRAS

By J. H. M. Wedderburn

(Received June 3, 1937)

1. The object of this note is to extend some of the methods which I have used elsewhere[1] so as to apply to algebras in modular fields.

Theorem 1. *If an algebra has a nilpotent basis, it is itself nilpotent.*

Let $u_1, u_2, \cdots, u_a$ be a basis of an algebra A which has an identity, each element of the basis being nilpotent. If N is the radical of A, we may suppose its basis forms a part of the given basis, and $A - N$ has then also a nilpotent basis; we therefore suppose that A is semi-simple. It is clear that each irreducible part of A also has a nilpotent basis since, if e_1 is the identity of one part and $e_2 = 1 - e_1$, then $u = e_1ue_1 + e_2ue_2 = v_1 + v_2$ and $u^r = v_1^r + v_2^r$. Hence it is sufficient to prove the theorem for a simple algebra.

Let $e_1, e_2, \cdots, e_n$ be a complete supplementary set of idempotent elements of A and set $A = \sum A_{ij}$, $A_{ij} = e_iAe_j$. If, say, A_{ii} has a radical N_i, then $AN_iA = Ae_iN_ie_iA$ does not contain e_i and, being therefore less than A, it forms an invariant subalgebra; but A is simple and hence A_{ii} has no radical and, since e_i is primitive, it is a division algebra. If we now extend the field, the basis is still composed of nilpotent elements and we may suppose the field so extended that each division algebra A_{ii} is of order 1.

Since $A_{ii} = e_i$, then $A_{ij}A_{ji}$ is either e_i or 0. If $A_{ij}A_{ji} = e_i$, we can find $e_{ij} < A_{ij}$, $e_{ji} < A_{ji}$ such that $e_{ij}e_{ji} = e_i$; for we cannot have $A_{ij}e_{ji} = 0$ for every e_{ji} and, if not 0, it equals e_i. Also $e_{ji}e_{ij}e_{ji} = e_{ji}e_i = e_{ji}$ and hence $e_{ji}e_{ij}$ is not 0 and therefore equals e_j. If we have also $e'_{ij}e_{ji} = e_i$, $e'_{ij} < A_{ij}$, then

$$0 = (e'_{ij} - e_{ij})e_{ji}e_{ij} = (e'_{ij} - e_{ij})e_{jj} = e'_{ij} - e_{ij},$$

and the same reasoning shows that we cannot have $e'_{ij}e_{ji} = 0$; hence for each $i \neq j$ for which $A_{ij}A_{ji} \neq 0$ we have $A_{ij} = e_{ij}$, $A_{ji} = e_{ji}$.

Suppose now that, say, $A_{12}A_{21} = 0$ from which it follows as above that $A_{21}A_{12} = 0$. Then if $N = A_{12} + A_{21}$ we have $e_1ANAe_1 = 0$ and, since A is simple, $N = 0$. Hence either $A_{ij} = 0$ or $A_{ij} = e_{ij}$ and, if we set $e_i = e_{ii}$, we have a basis for A which forms a part of the ordinary matric basis for matrices of order n. Since in the matric algebra the trace of e_{11} is 1, A cannot have a nilpotent basis and hence the theorem is proved.[2]

2. It is convenient to repeat here with some amplifications a theorem which was given many years ago.[3]

---

[1] *Lectures on Matrices*, New York, 1934, pp. 155, 159, quoted hereafter as "Lectures".

[2] It is easily seen that no e_{ij} is 0 but this is not required in the proof.

[3] Wedderburn, *Hypercomplex Numbers*, Proc. Lond. Math. Soc. (2) 6, 1907, p. 105. Inseparable extensions were not considered in this paper.

THEOREM 2. *If A is a commutative algebra whose identity is its only primitive idempotent element, it can be expressed in the form $A = D + N$ where N is the radical and D is a field isomorphic with $A - N$, unless the derivative of the characteristic function of $A - N$ is identically zero.*

Since $A - N$ is a field, its characteristic function $f(\lambda)$ is necessarily of the same degree as the order of the field when $f'(\lambda) \not\equiv 0$ and, when this condition is satisfied, there is a primitive element x' whose characteristic function $g(\lambda)$ has the same degree as $f(\lambda)$ and has no multiple roots. If x is an element of A corresponding to x' in $A - N$, then, since $g(x') = 0$, $g(x)$ lies in N and hence there is a minimal integer r such that $[g(x)]^r = 0$. We now seek a rational polynomial in x, say z, such that $g(z) = 0$, $z = x$ (mod. N). Let $z = x + u$; then[4]

$$0 = g(z) = g(x) + g'(x)u + \frac{g''(x)}{2}u^2 + \cdots + \frac{g^{(r-1)}(x)}{(r-1)!}u^{r-1}. \tag{1}$$

Under the given conditions the coefficient of u has an inverse and it is readily proved by induction that (1) can be formally inverted as a series of the form

$$u = -\frac{g(x)}{g'(x)} + \cdots = h_1(x)g(x) + h_2(x)[g(x)]^2 + \cdots + h_{r-1}(x)[g(x)]^{r-1} \tag{2}$$

where the h's are rational functions of x with rational coefficients whose denominators are not singular and where the series terminates since $[g(x)]^r = 0$. The algebra generated by z is then of the same order as $A - N$ and the theorem is therefore proved.

The following example shows that this result does not necessarily follow when $f'(\lambda) = 0$. Let F be the field obtained by adjoining an indeterminate t to the field of rational integers reduced modulo 2, and let A be the algebra generated by an element a which satisfies the equation $a^4 + t^2 = 0$. Here $N = (a^2 + t, a(a^2 + t))$ and, if we set $A = (1, b, a^2 + t, a(a^2 + t))$, we have

$$b = \beta_0 + \beta_1 a + \beta_2(a^2 + t) + \beta_3 a(a^2 + t)$$
$$b^2 = \beta_0^2 + \beta_1^2 a^2 = \beta_0^2 + \beta_1^2 t + \beta_1^2(a^2 + t).$$

If b^2 is linearly dependent on $(1, b)$, we must have $\beta_1 = 0$, which is impossible since $(1, b, a^2 + t, a(a^2 + t))$ is a basis of A. Hence in this case the field cannot be separated from the radical.

It should be noticed that the algebra, $A - N$, defined by $a^2 + t = 0$ is not semi-simple when the field is extended by the adjuntion of $\beta = t^{\frac{1}{2}}$; for then $(a - \beta)^2 = 0$ and, if $b = a - \beta$, the extended algebra has the basis $(1, b)$ in which $b^2 = 0$.

3. THEOREM 3. *If the field F of a semi-simple algebra A is extended by the adjunction of a root of a separable equation, the algebra remains semi-simple.*

---

[4] The coefficient of u^s is determinate even if $s \equiv 0$ in the field.

We may suppose without loss of generality that A is simple, say $A = D \times M$, where $M = (e_{pq})$ is a simple matric algebra and D a division algebra. If after the adjunction of an irrationality ξ of degree $m + 1$, A is no longer semi-simple, it will contain a radical N with an element of the form

$$n = a_0 + a_1 \xi + \cdots + a_m \xi^m \tag{3}$$

where $a_0, a_1, \cdots$, have coefficients in the original field when expressed in terms of the original basis. We may suppose that the a's are in D; for $\sum_i e_{ip} a e_{qi}$ is in D and $\sum_i e_{ip} n e_{qi}$ in N. If the a's are not commutative, say $a_i a_j \neq a_j a_i$, then $a_i n - n a_i$ is in N, is not 0, and contains fewer terms than (3). Hence by a repetition of this process we obtain an element of the radical of the form (3) in which the a's are commutative. The a's then generate a commutative division algebra $F(a_0, a_1, \cdots)$ which can also be regarded as a field over F. Let the irreducible equation in F of which ξ is a root be $f(\lambda) = 0$ and set $h(\lambda) = a_0 + a_1 \lambda + \cdots + a_m \lambda^m$. If $k(\lambda)$ is the H.C.F. of $f(\lambda)$ and $h(\lambda)$, we can find polynomials $\alpha(\lambda)$, $\beta(\lambda)$ with coefficients in $F(a_0, a_1, \cdots)$ such that

$$\alpha(\lambda) h(\lambda) + \beta(\lambda) f(\lambda) \equiv k(\lambda).$$

This gives $\alpha(\xi) h(\xi) = k(\xi)$ and hence $k(\xi)$ lies in N. Let $f(\lambda) = k(\lambda) k_1(\lambda)$; since $k(\lambda)$ and $k_1(\lambda)$ can have no factor in common when $f(\lambda)$ is separable, there exist polynomials $\delta(\lambda)$, $\gamma(\lambda)$ such that

$$\delta(\lambda) k(\lambda) + \gamma(\lambda) k_1(\lambda) \equiv 1,$$

and, if $n_1 = \delta(\xi) k(\xi) < N$, $b = \gamma(\xi) k_1(\xi)$, we have $n_1 + b = 1$ so that b has the inverse $1 + n_1 + n_1^2 + \cdots$. Hence $k_1(\xi)$ has an inverse. But $k(\xi) k_1(\xi) = f(\xi) = 0$ so that we must have $k(\xi) = 0$, from which it follows that $h(\xi) = 0$ so that the radical is 0. The theorem is therefore proved.

The preceding theorems are sufficient for the discussion of the classification of algebras given in my Lectures with one apparent exception on page 163 where the number 2 appears in the inverse of an element, but this is easily avoided by replacing u by $1 + ae$ and v by $1 - ea$.

PRINCETON, N. J.

Annals of Mathematics
Vol. 38, No. 4, October, 1937

ON ALGEBRAS WHICH ARE CONNECTED WITH THE SEMISIMPLE CONTINUOUS GROUPS*

By Richard Brauer

(Received March 5, 1937)

1. **Introduction.** We consider a group $\mathfrak{G}$ of linear transformations in an n-dimensional vector space V_n. If a transformation G of $\mathfrak{G}$ is performed, the components of a general tensor of rank f undergo a linear transformation $M_f(G)$ and these $M_f(G)$ form a representation $\mathfrak{M}_f$ of $\mathfrak{G}$. The investigation of $\mathfrak{M}_f$ is of great importance for the theory of representations. In particular, we have to study the breaking up of $\mathfrak{M}_f$ into its irreducible constituents. When dealing with this question we may replace $\mathfrak{M}_f$ by its enveloping algebra A_f, i.e. the totality of all matrices which can be written as linear combinations of matrices of $\mathfrak{M}_f$ with scalar coefficients.

Here at once the problem arises of giving a direct characterization of the matrices belonging to this algebra A_f. If $\mathfrak{G}$ is the full linear group consisting of all non-singular linear transformations, we can easily give the answer. The matrices of A_f can be characterized by certain conditions of symmetry. There is no essential difference between the case of this group $\mathfrak{G}$ and the case of the unimodular group where one considers only the transformations of determinant 1.

However, one meets with difficulties in the case of the other semisimple continuous groups. H. Weyl[1] determined the enveloping algebra A_f, if $\mathfrak{G}$ is the complex group. But in his proof he had to make use of the results of the theory of representations for this group $\mathfrak{G}$. It seems desirable to give a method which is independent of this theory, in order to be able to develop the theory of representations starting from the investigation of the algebra A_f. In the case of the orthogonal group the direct characterization of A_f has not yet been given.

It is my aim to give an elementary method which works for all semisimple groups. Instead of A_f we first consider the commutating algebra B_f, which consists of all matrices commutative with every matrix of A_f. In §2 we treat for any group $\mathfrak{G}$ the problem of finding a basis of B_f, and show that it is equivalent to the problem answered by the fundamental theorem of invariant theory. We may characterize A_f as the commutating algebra of B_f, provided we know that A_f is semisimple, or in other words that the representation $\mathfrak{M}_f$ of $\mathfrak{G}$ is completely reducible. This follows for a semisimple group from the general theory of representation, if the underlying field is the field of all real or all complex numbers. In §3 a direct proof of the complete reducibility will be

---

* Presented to the American Mathematical Society, September 1, 1936.

[1] Mathematische Zeitschrift *35*, p. 300 (1932).

given if $\mathfrak{G}$ is the complex group or the orthogonal group, and **K** is any field of characteristic 0. Thereby we obtain a characterization of the elements of A_f in these cases. In the case of the full linear group or that of the unimodular group $\mathfrak{G}$ we use our method in the other direction (§4). Here the algebra A_f is known from the beginning, and we can derive the algebra B_f from it. The results of §2 then give a proof of the fundamental theorem of invariant theory.

In §5 we study the algebra B_f more closely. If $\mathfrak{G}$ is the full linear group, it is well known that B_f is homomorphic to the group ring of the symmetric group $\mathfrak{S}_f$ of permutations on f symbols. This is the basis for the important connection between the representations of the full linear group and the symmetric permutation group.[2] It seems, therefore, that the algebras which play the same rôle for the other simple continuous groups are of some interest.

In a later paper, I shall investigate the algebras A_f and B_f further and apply these considerations to the theory of representations and to a proof of the second fundamental theorem of invariant theory.

2. **The commutating algebra. Equivalence with the fundamental theorem of invariant theory.** We consider a group $\mathfrak{G}$ of linear transformations

$$x_i' = \sum_k g_{ik} x_k \tag{1}$$

in an n-dimensional vector space V_n with the coordinates $x_1, x_2, \cdots, x_n$. The coefficients g_{ik} of G and the x_i may here belong to any given field **K**.

Let $t_{i_1 i_2 \cdots i_f}$ be the components of a tensor of rank f, ($i_\rho = 1, 2, \cdots, n$), for a fixed f. These components undergo a linear transformation induced by the transformation (1)

$$t'_{i_1 i_2 \cdots i_f} = \sum_{k_1, k_2, \cdots, k_f} g_{i_1 k_1} g_{i_2 k_2} \cdots g_{i_f k_f} t_{k_1 k_2 \cdots k_f}. \tag{2}$$

We denote the system $i_1, i_2, \cdots, i_f$ of f indices by $[i]$ and similarly the system $k_1, k_2, \cdots, k_f$ by $[k]$, and write (2) in the form

$$t([i])' = \sum_{[k]} m([i], [k]) t([k]) \tag{3}$$

where we have set

$$\begin{aligned} t([i]) &= t_{i_1 i_2 \cdots i_f}, \\ m([i], [k]) &= g_{i_1 k_1} g_{i_2 k_2} \cdots g_{i_f k_f}. \end{aligned} \tag{4}$$

In the sum (3), $[k]$ is ranging over all n^f systems $[k]$, ($k_1, k_2, \cdots, k_f = 1, 2, \cdots, n$). We arrange these n^f systems $[k]$ in a fixed order (for instance, choosing

---

[2] This connection has first been given by I. Schur, *Dissertation*, Berlin, 1901. Cf. in addition I. Schur, Sitzungsberichte d. Preuss. Akad. 1927, p. 58, H. Weyl, *The Theory of Groups and Quantum Mechanics*, London, 1932, Chapter 5, and also van der Waerden, Math. Ann. *104*, p. 92 and p. 800, 1931. The method given in this last paper is essentially the same as Schur's second method.

the lexicographical arrangement). In the matrix $M(G)$ of the transformation (3), a row is given by a symbol $[i]$ and a column similarly by a symbol $[k]$. In the row $[i]$ and the column $[k]$, we have the coefficient (4).

$$M(G) = (m([i], [k])). \tag{5}$$

If two transformations G and H of $\mathfrak{G}$ are performed one after the other, the tensor $t_{i_1 i_2 \cdots i_f}$ undergoes first $M(G)$ and then $M(H)$. We readily obtain $M(HG) = M(H)M(G)$. This shows that the matrices $M(G)$ form a representation $\mathfrak{M}$ of $\mathfrak{G}$.

We denote by A the system consisting of all linear combinations, $A = c_1 M_1 + c_2 M_2 + \cdots + c_r M_r$, of any finite system $M_1, M_2, \cdots, M_r$ of elements of $\mathfrak{M}$ with coefficients c_ρ in **K**. Obviously, A forms an algebra, *the enveloping algebra* of $\mathfrak{M}$. We write the matrices A in the same manner as $M(G)$ in (5),

$$A = (a([i], [k])) \tag{6}$$

where again the symbol $[i]$ gives the row and $[k]$ gives the column. Our problem is to find a system of relations for the $(n^f)^2$ coefficients $a([i], [k])$ which form the necessary and sufficient conditions that a matrix A, (6), actually belongs to A.

At the same time with A we consider the matrices

$$B = (b([i], [k])) \tag{7}$$

of the same degree n^f which are commutative with every element A of A. These matrices B also form an algebra B, the *commutating algebra* of A.

It is, of course, sufficient to require that B is commutative with $M(G)$ for every G in $\mathfrak{G}$,

$$M(G)\, B = BM(G). \tag{8}$$

This gives

$$\sum_{[j]} m([i], [j]) b([j], [k]) = \sum_{[j]} b([i], [j]) m([j], [k]). \tag{9}$$

We consider now f vectors $\mathfrak{x}(1), \mathfrak{x}(2), \cdots, \mathfrak{x}(f)$,

$$\mathfrak{x}(\rho) = (x_1(\rho), x_2(\rho), \cdots, x_n(\rho)) \tag{10}$$

which undergo cownerrediently the transformation G of $\mathfrak{G}$ and, furthermore, vectors $\mathfrak{u}(1), \mathfrak{u}(2), \cdots, \mathfrak{u}(f)$,

$$\mathfrak{u}(\rho) = (u^{(1)}(\rho), u^{(2)}(\rho), \cdots, u^{(n)}(\rho)) \tag{11}$$

which undergo the contragredient transformation G'^{-1},

$$x_j(\rho)' = \sum_k g_{jk} x_k(\rho), \tag{12}$$

$$u^{(j)}(\sigma) = \sum_i g_{ij} u^{(i)}(\sigma)'. \tag{13}$$

We then have (cf. (4)),

$$(14)\qquad x_{j_1}(1)'x_{j_2}(2)' \cdots x_{j_f}(f)' = \sum_{[k]} m([j], [k])x_{k_1}(1)x_{k_2}(2) \cdots x_{k_f}(f),$$

$$(15)\qquad u^{(j_1)}(1)u_i^{(j_2)}(2) \cdots u^{(j_f)}(f) = \sum_{[i]} m([i], [j])u^{(i_1)}(1)'u^{(i_2)}(2)' \cdots u^{(i_f)}(f)'.$$

We multiply (9) by

$$(16)\qquad x_{k_1}(1)x_{k_2}(2) \cdots x_{k_f}(f)u^{(i_1)}(1)'u^{(i_2)}(2)' \cdots u^{(i_f)}(f)'$$

and add over all $[i]$, $[k]$. On the right-hand side, we use (14) and obtain

$$(17)\qquad \sum_{[i],[j]} b([i], [j])u^{(i_1)}(1)' \cdots u^{(i_f)}(f)'x_{j_1}(1)' \cdots x_{j_f}(f)'.$$

On the left-hand side, we use (15) and obtain

$$(18)\qquad \sum_{[j],[k]} b([j], [k])u^{(j_1)}(1) \cdots u^{(j_f)}(f)x_{k_1}(1) \cdots x_{k_f}(f).$$

Comparison of (17) and (18) shows that

$$(19)\qquad J = \sum_{[i],[k]} b([i], [k])u^{(i_1)}(1) \cdots u^{(i_f)}(f)x_{k_1}(1) \cdots x_{k_f}(f)$$

is an invariant for the group $\mathfrak{G}$.[3]

If, conversely, an invariant J of $\mathfrak{G}$ is given which depends linearly and homogeneously on f cogredient vectors (10) and f contragredient vectors (11), we may set J in the form (19). We may now retrace each step: The invariance of J shows that (17) and (18) are equal for every G in $\mathfrak{G}$. Using (14) and (15) and comparing the coefficients of (16) on both sides we come back to (8).

This shows that *the coefficients $b([i], [k])$ of the matrices B of* B *are characterized by the fact that they appear as the coefficients of the invariants J*, (19).

Let $J_1, J_2, \cdots, J_N$ be such a system of invariants (19) that every invariant J of this type can be written in the form $J = c_1J_1 + c_2J_2 + \cdots + c_NJ_N$, with constant coefficients c_ρ in K. If J_ρ corresponds to the matrix B_ρ in B and J to the matrix B, we have

$$(20)\qquad B = c_1B_1 + c_2B_2 + \cdots + c_NB_N.$$

The linearly independent matrices among $B_1, B_2, \cdots, B_N$ constitute a basis of B.

In the cases of the groups $\mathfrak{G}$ which are of principal interest for us, the main theorem of invariant theory allows the construction of a system $J_1, J_2, \cdots, J_N$. The corresponding elements $B_1, B_2, \cdots, B_N$ in B yield a construction of the most general element B of B in the form (20). It is, in general, not important for our purpose to select a linearly independent basis among $B_1, B_2, \cdots, B_N$.

[3] An alternative proof is obtained as follows: One may write (8) in the form $B = M(G)BM(G^{-1})$. This equation shows that $b([i], [k])$ is an invariant tensor under the transformation (1), if $i_1, i_2, \cdots, i_f$ are treated as covariant indices and $k_1, k_2, \cdots, k_f$ as contragredient indices. Therefore, the $b([i], [k])$ are the coefficients of an invariant, depending linearly on f contragredient vectors and f cogredient vectors.

The commutating algebra A^* of B certainly contains A, $\mathrm{A}^* \geqq \mathrm{A}$. Moreover, if we know that A is semisimple, the theorem holds[4] that B also is semisimple, and that A itself is the commutating algebra of B. *In this case, the elements A of A may be characterized by the equations*

$$AB_\rho = B_\rho A, \qquad (\rho = 1, 2, \cdots, N) \tag{21}$$

and this gives a solution of the problem formulated in §1.

The fact that A is semisimple is equivalent to the fact that A, as a system of matrices, is completely reducible, and this again is equivalent to the complete reducibility of $\mathfrak{M}$. If K is the field of all real or of all complex numbers, and if $\mathfrak{G}$ is a semisimple continuous group, it is shown in the theory of representations that every representation $\mathfrak{M}$ is completely reducible. Therefore, the method given here can be applied in these cases. We shall, however, prove the complete reducibility for the most important of these groups in §§3, 4 by direct elementary methods.

The result of this §2 can be slightly generalised. If we denote the representation $M(G)$ of $\mathfrak{G}$ belonging to the tensors of rank f by $M_f(G)$, we may ask which matrices B satisfy the relation $M_f(G)\ B = BM_{f^*}(G)$ for all G in $\mathfrak{G}$ and fixed f and f^*. We shall put here $B = (b([i], \cdot \{k\}))$, where $[i]$ has the same significance as above whereas $\{k\}$ ranges over all systems $k_1, k_2, \cdots, k_{f^*}$ of f^* indices k_ρ, $(k_\rho = 1, 2, \cdots, n)$. The same method applies. The only difference is that J here will be an invariant depending on f contragredient vectors $\mathfrak{u}(1)$, $\mathfrak{u}(2), \cdots, \mathfrak{u}(f)$ and f^* cogredient vectors $\mathfrak{x}(1), \cdots, \mathfrak{x}(f^*)$.

3. The complete reducibility for the case of the orthogonal group and the complex group.

a) *The orthogonal group.* Let $\mathfrak{G}$ be the group of all orthogonal transformations, first for the case when $\mathrm{K} = \mathrm{P}$ is the field of all rational numbers. Then $\mathfrak{M}$ also consists of orthogonal matrices and is, therefore, completely reducible. Hence A in this case is semisimple. We choose a basis

$$A_1, A_2, \cdots, A_r \tag{22}$$

in A and a basis

$$B_1, B_2, \cdots, B_s \tag{23}$$

in B. Every matrix commutative with all the A_ρ lies in B, every matrix commutative with all the B_σ lies in A. There are invariants $J_1, J_2, \cdots, J_s$ corresponding to $B_1, B_2, \cdots, B_s$. Here J_σ is invariant under all orthogonal transformations $\tilde{G}$ with real coefficients because every such $\tilde{G}$ is the limit of orthogonal transformations G with rational coefficients.[5] It follows from the main theorem of invariant theory that J_σ is a polynomial in the inner products

[4] Cf., for instance, H. Weyl, Annals of Math (2), 37, p. 709 (1936), theorem (1.4 – B).

[5] This fact easily follows from a parametric representation of the orthogonal group.

of any two of the vectors, and, if we restrict ourselves to the group of proper orthogonal transformations (determinant $+1$), in the determinants of n of the vectors. We replace now $\mathrm{K} = \mathrm{P}$ by any field K of characteristic 0. The last consideration shows that J_σ still is an invariant of $\mathfrak{G}$ and, therefore, B_σ still belongs to B. On the other hand, it is trivial that A_ρ still belongs to A.

Each of the systems (22) and (23) still constitutes a maximum system of linear independent matrices commutative with every matrix of the other system. This shows that (22) still is a basis of A and (23) is a basis of B. It follows that A is completely reducible, and this is what we wanted to prove. At the same time, we readily see that the fundamental theorem of invariant theory holds for any field K of characteristic 0.

We can now easily give the invariants J_ρ and the matrices B_ρ (§2), and set up the equations (21) in explicit form, but we will postpone this until §5.

b) *The complex group.* Let $\mathfrak{G}$ now be the complex-group with coefficients in any field K of characteristic 0. At the end of this section we shall prove

LEMMA 1. *Let $f(g_{11}, g_{12}, \cdots, g_{nn}) = f(g_{ik})$ be a polynomial in n^2 variables. If f vanishes for all systems g_{ik} for which the matrix $(g_{ik}) = G$ belongs to $\mathfrak{G}$ and all the g_{ik} are rational, then f vanishes for all g_{ik} for which the matrix $G = (g_{ik})$ belongs to $\mathfrak{G}$ and the g_{ik} lie in* K.

The coefficients $m([i], [k])$ of $M(G)$ are polynomials $f(g_{ik})$ and so are the coefficients in any equivalent representation $P^{-1}M(G)P$, where P is a fixed non-singular matrix.

If we have complete reducibility of the representation $\mathfrak{M}$ in the case that $\mathrm{K} = \mathrm{P}$ is the field of all rational numbers, the lemma shows that we have complete reducibility for any other K of characteristic 0. If in the case $\mathrm{K} = \mathrm{P}$ complete reducibility of $\mathfrak{M}$ does not hold, it does not hold for any other K of characteristic 0. It is, therefore, sufficient to investigate the question of complete reducibility for a fixed field K of characteristic 0, so we choose the field of all complex numbers. We are going to show that $\mathfrak{M}$ actually is completely reducible. We use a second lemma which we will prove simultaneously with lemma 1.

LEMMA 2. *If a polynomial $f(g_{ik})$ vanishes for all g_{ik} for which $G = (g_{ik})$ lies in $\mathfrak{G}$ and is unitary, then $f(g_{ik})$ vanishes for all $(g_{ik}) = G$ in $\mathfrak{G}$.*

We shall assume that lemma 2 has been proved. We denote by $\mathfrak{G}^*$ the subgroup of all elements of $\mathfrak{G}$ which are unitary. Let $\mathfrak{M}^*$ be the representation of $\mathfrak{G}$ formed by the matrices $M(G^*)$ for G^* in $\mathfrak{G}^*$. Since G^* is unitary, so is $M(G^*)$ and, therefore, $\mathfrak{M}^*$ is completely reducible. We may find a matrix P with constant coefficients such that $P^{-1}M(G^*)P$ breaks up completely into irreducible constituents. The lemma then shows that $P^{-1}M(G)P$, G in $\mathfrak{G}$, breaks up in exactly the same way, and the constituents here are irreducible because they are so when we restrict our attention to the subgroup $\mathfrak{G}^*$. Hence $\mathfrak{M}$ also is completely reducible.

PROOF OF LEMMA 1 AND LEMMA 2. We denote the variables in the space

V_n here by $x_1, x_{1'}, x_2, x_{2'}, \cdots, x_\nu, x_{\nu'}$; $n = 2\nu$, and take the invariant skew-symmetric form in the form

$$\{\mathfrak{x}, \mathfrak{y}\} = \sum_{p=1}^{\nu} (x_p y_{p'} - x_{p'} y_p) = \sum_{i,k=1,1',\cdots,\nu,\nu'} \epsilon(i, k) x_i y_k. \tag{24}$$

If $Z = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a unimodular matrix of degree 2, $ad - bc = 1$, the transformation $P_i(Z)$,

$$\begin{aligned} x_i' &= ax_i + bx_{i'} \\ x_{i'}' &= cx_i + dx_{i'} \\ x_l' &= x_l \qquad \text{for } l \neq i, i'; l = 1, 1', 2, 2', \cdots, \nu, \nu', \end{aligned}$$

belongs to $\mathfrak{G}$ and so does the transformation $Q_{i,j}(Z)$,

$$\begin{aligned} x_i' &= ax_i \qquad + bx_j \\ x_{i'}' &= \qquad dx_{i'} \qquad - cx_{j'} \\ x_j' &= cx_i \qquad + dx_j \\ x_{j'}' &= \qquad -bx_{i'} \qquad + ax_{j'} \\ x_l' &= \qquad x_l \qquad \text{for } l \neq i, i', j, j'. \end{aligned}$$

One sees easily[6] that there exists a product

$$R_1(Z_1)R_2(Z_2) \cdots R_t(Z_t) = \Pi \tag{25}$$

with the following properties: (1) Every $R_\tau(Z_\tau)$ is either a $P_i(Z_i)$ or a $Q_{ij}(Z_i)$. (2) If G is an element of $\mathfrak{G}$ the unimodular matrices Z_i can be chosen in such a way that $G\Pi$ has the form

$$G\Pi = \left(\begin{array}{cc|c} 1 & 0 & 0\cdots0 \\ 0 & 1 & 0\cdots0 \\ \hline \multicolumn{2}{c|}{T} & U \end{array}\right) \tag{26}$$

The matrix $G\Pi$, (26), lies in $\mathfrak{G}$. Therefore, we have 0 at the places indicated by T. The matrix U will belong to the complex-group for the case of n-2 dimensions. We deal with U in a similar manner as with G now only taking P_i with $i > 1$ and Q_{ij} with $i > 1$, $j > 1$. We finally obtain a product Π of the same form as (25) such that $G\Pi$ equals the unit matrix I. Then $G = \Pi^{-1}$. Since the inverse of every $R_\tau(Z_\tau)$ is of the same form, we obtain a product

$$S_1(Z_1)S_2(Z_2) \cdots S_l(Z_l) \tag{27}$$

---

[6] The argument is similar to the one used by E. Mohr, *Dissertation*, Göttingen, 1933, but for our purpose, Mohr's generating elements are not suitable.

such that every S_τ is either a P_i or a Q_{ij}, (2) every G in $\mathfrak{G}$ can be obtained in (27) if the Z_i are properly chosen.

Conversely, it is clear that (27) belongs to $\mathfrak{G}$. If all the Z_i have rational coefficients, so has (27). Further, if all the Z_i are unitary, so is (27). From (27) we derive a parametric representation of $\mathfrak{G}$

$$G = G(u_1, u_2, \cdots, u_q) \tag{28}$$

such that (28) lies in $\mathfrak{G}$ for all values of $u_1, u_2, \cdots, u_q$, and every element of $\mathfrak{G}$ can be obtained in this form[7], and that, furthermore, G has rational coefficients if all $u_1, u_2, \cdots, u_q$ are rational.

We express now, in the case of lemma 1, all the g_{ik} in $f(g_{ik})$ by $u_1, u_2, \cdots, u_q$. We obtain then a polynomial $F(u_1, u_2, \cdots, u_q)$ which vanishes if $u_1, u_2, \cdots, u_q$ are rational. Then F vanishes identically in $u_1, u_2, \cdots, u_q$ and this gives lemma 1.

Secondly, we may derive from (27) a similar parametric representation

$$G = G^*(v_1, v_2, \cdots, v_r)$$

such that G lies in $\mathfrak{G}$ for all $v_1, v_2, \cdots, v_r$ and every element of $\mathfrak{G}$ can be obtained in this form, and that, furthermore, G is unitary if all the v_ρ are real. An argument similar to that used in the case of lemma 1, gives lemma 2.

4. **Proof of the fundamental theorem of invariant theory for the full linear group.** We consider now the full linear group $\mathfrak{G}$. Let $[i]$, as in §2, be a system of f indices $(i_1, i_2, \cdots, i_f)$; $(i_\rho = 1, 2, \cdots, n)$. If

$$P = \begin{pmatrix} 1 & 2 & \cdots & f \\ p_1 & p_2 & \cdots & p_f \end{pmatrix} \tag{29}$$

is a permutation of $1, 2, \cdots, f$, we denote by $[i]_P$ the system $(i_{p_1}, i_{p_2}, \cdots, i_{p_f})$. From (4) or

$$m([i], [k]) = g_{i_1k_1} g_{i_2k_2} \cdots g_{i_fk_f},$$

it follows immediately that we have

$$m([i]_P, [k]_P) = m([i], [k]).$$

Of course, the corresponding relations

$$a([i]_P, [k]_P) = a([i], [k]) \tag{30}$$

hold for every element $A = (a([i], [k])$ of the enveloping algebra **A**.

We construct a set of systems $([i], [k])$,

$$([i^{(1)}], [k^{(1)}]), ([i^{(2)}], [k^{(2)}]), \cdots, ([i^{(M)}], [k^{(M)}])$$

such that every $([i], [k])$ can be obtained in the form $([i^{(\mu)}]_P, [k^{(\mu)}]_P)$ for a properly chosen permutation P and for exactly one μ. The $m([i^{(\mu)}], [k^{(\mu)}])$ are then all different for $\mu = 1, 2, \cdots, M$. A linear homogeneous relation between the

[7] We need not consider exceptional elements, if we take q large enough and use supernumerary parameters.

$m([i^{(\mu)}], [k^{(\mu)}])$ with constant coefficients c_μ holds for independent g_{ik} if it holds for all g_{ik} with non-vanishing determinant. If follows that in such a relation all coefficients c_μ vanish. This shows that $\mathfrak{M}$ and hence A contains exactly M linearly independent elements and that the coefficients $a([i^{(\mu)}], [k^{(\mu)}])$, $(\mu = 1, 2, \cdots, M)$ can be arbitrarily assigned in an element of A. Consequently, that the equations (30) be satisfied is a necessary and sufficient condition that a matrix A belongs to A.[8]

We set $\delta(i, k) = 0$ for $i \neq k$, $\delta(i, k) = 1$ for $i = k$ and form the matrix

$$B_P = (\delta(i_1, k_{p_1})\, \delta(i_2, k_{p_2}) \cdots \delta(i_f, k_{p_f})) = (\delta([i], [k]_P) \tag{31}$$

for every permutation (29). We find then for any matrix $C = (c([i], [k]))$ of degree n^f the relation

$$CB_P = \Big(\sum_{[j]} c([i], [j])\, \delta([j], [k]_P)\Big) = (c([i], [k]_P)).$$

In particular, for two permutations P and Q of $1, 2, \cdots, f$, we have

$$B_P B_Q = B_{PQ},$$

and the B_P form a representation of the symmetric group $\mathfrak{S}_f$, consisting of all permutations of f indices $1, 2, \cdots, f$. Moreover, we have

$$B_P^{-1}C = B_{P-1}C = \Big(\sum_{[j]} \delta([i], [j]_{P-1})c([j], [k])\Big) = (c([i]_P, [k])).$$

Hence

$$B_P^{-1}CB_P = (c([i]_P, [k]_P)).$$

The equations (30) therefore express the fact that A is commutative with all B_P. Let us denote by B* the enveloping algebra of the representation B_P of $\mathfrak{S}_f$. Then A is the commutating algebra of B*. As a representation of a finite group the representation B_P of $\mathfrak{S}_f$ is completely reducible provided that the underlying field K has the characteristic 0. It follows[9] that A is completely reducible and that B*, conversely, is the commutating algebra of A. Hence B* = B.

Every element of B can be written as a linear combination of the special elements B_P. It follows that every invariant J, (19), is a linear combination of the invariants J_P corresponding to the B_P. According to (7), (31) and (19) we have to set

$$\begin{aligned} J_P &= \sum_{[i],[k]} \delta(i_1, k_{p_1}) \cdots \delta(i_f, k_{p_f}) u^{(i_1)}(1) \cdots u^{(i_f)}(f) x_{k_1}(1) \cdots x_{k_f}(f) \\ &= \sum_{[k]} u^{(k_{p_1})}(1) x_{k_1}(1) \cdots u^{(k_{p_f})}(f) x_{k_f}(f) \\ &= \sum_{j_1} u^{(j_1)}(1) x_{j_1}(p_1) \cdots \sum_{j_f} u^{(j_f)}(f) x_{j_f}(p_f), \end{aligned}$$

$$J_P = (\mathfrak{u}(1), \mathfrak{x}(p_1))(\mathfrak{u}(2), \mathfrak{x}(p_2)) \cdots (\mathfrak{u}(f), \mathfrak{x}(p_f)) \tag{32}$$

---

[8] Cf. H. Weyl, Annals of Math. (2) *30*, p. 499, (1929) and van der Waerden[2].

[9] Cf.[4].

where $(\mathfrak{u}, \mathfrak{x}) = \sum u^{(i)} x_i$ denotes the inner product of a cogredient vector $\mathfrak{x}$ and a contragredient vector $\mathfrak{u}$. The fact that every invariant J, (19), can be written as a linear combination of the invariants (32) gives the special case of the fundamental theorem of invariant theory in which we have the same number of cogredient and contragredient vectors.

It is not difficult to derive the general theorem from the special case treated here. I give this argument for the sake of completeness. Let J be an invariant depending linearly and homogeneously on f cogredient vectors $\mathfrak{x}(1), \mathfrak{x}(2), \cdots, \mathfrak{x}(f)$ and h contragredient vectors $\mathfrak{u}(1)\ \mathfrak{u}(2), \cdots, \mathfrak{u}(h)$. We have

$$J' = (\det g_{ik})^r \cdot J \tag{33}$$

where J' is the value of the invariant for the transformed vectors $\mathfrak{x}(i)'\ \mathfrak{u}(i)'$, cf. (12), (13). On comparing the degree with regard to all the g_{ik} in (33), we have

$$f = nr + h. \tag{34}$$

The fundamental theorem of invariant theory states that J is a polynomial Z in the products $(\mathfrak{u}(i), \mathfrak{x}(k))$, the determinants of n of the vectors $\mathfrak{x}(i)$, and the determinants of n of the vectors $\mathfrak{u}(i)$. The determinants of the first kind occur only if $f > h$, and every term in Z contains r such determinants. The determinants of the second kind occur only if $f < h$, and every term in Z contains $|r|$ such determinants. Let us assume that the theorem has already been proved for smaller values of $|r|$; it is true for $r = 0$ as we have seen. We may assume that $f > h$, otherwise we interchange the $\mathfrak{x}(i)$ and the $\mathfrak{u}(i)$. We introduce n new contragredient vectors $\mathfrak{u}(h + 1), \cdots, \mathfrak{u}(h + n)$. Let D be their determinant. Then DJ is an invariant which, according to (34), has the weight $r - 1$. Therefore, for DJ the theorem is true. Hence, we obtain a representation

$$DJ = \sum_{p_1 p_2 \cdots p_n} (\mathfrak{u}(h+1), \mathfrak{x}(p_1)) \cdots (\mathfrak{u}(h+n), \mathfrak{x}(p_n)) Z_{p_1 p_2 \cdots p_n}, \tag{35}$$

where $Z_{p_1 p_2 \cdots p_n}$ is a polynomial in the products $(\mathfrak{u}(i), \mathfrak{x}(k))$, $(i \leqq h)$ and the determinants of n vectors $\mathfrak{x}(i)$, every term containing $r - 1$ such determinants. The equation (35) holds for every system $\mathfrak{u}(h + 1), \cdots, \mathfrak{u}(h + n)$ of n vectors. We may, therefore, compare the coefficients of every term $u^{(i_1)}(h + 1) u^{(i_2)}(h + 2) \cdots u^{(i_n)}(h + n)$ on both sides. We thus obtain

$$\chi(i_1, i_2, \cdots, i_n) \cdot J = \sum_{p_1, p_2, \cdots, p_n} x_{i_1}(p_1) \cdots x_{i_n}(p_n) Z_{p_1 p_2 \cdots p_n} \tag{36}$$

where

$$\chi(i_1, i_2, \cdots, i_n) = \begin{cases} 0, \text{ if two of the } i_\rho \text{ are equal} \\ 1, \text{ if } i_1, i_2, \cdots, i_n \text{ form an even permutation} \\ \qquad \text{of } 1, 2, \cdots, n \\ -1, \text{ if } i_1, i_2, \cdots, i_n \text{ form an odd permutation} \\ \qquad \text{of } 1, 2, \cdots, n. \end{cases} \tag{37}$$

On multiplying (36) by $\chi(i_1, i_2, \cdots, i_n)$ and adding over all $i_1, i_2, \cdots, i_n$ $(i_\nu = 1, 2, \cdots, n)$, we obtain

$$n!J = \sum_{p_1, p_2, \cdots, p_n} | \mathfrak{x}(p_1), \mathfrak{x}(p_2), \cdots, \mathfrak{x}(p_n) | Z_{p_1 p_2 \cdots p_n}$$

where the factor of $Z_{p_1, p_2, \cdots, p_n}$ on the right is the determinant of $\mathfrak{x}(p_1), \mathfrak{x}(p_2), \cdots,$ $\mathfrak{x}(p_n)$. This proves our statement.

If K is a field of characteristic $p \neq 0$, our consideration can be applied as long as $f < p$, $h < p$. Every representation of the symmetric group $\mathfrak{S}_f$ still is completely reducible for its order $f!$ is not divisible by p. In particular, the representation $\mathfrak{M}_f$ is completely reducible for $f < p$, and it splits into irreducible parts in exactly the same way as in the case of characteristic 0.

5. **Construction of the algebras A and B.**

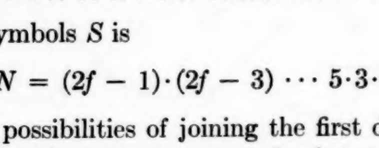
a) *The orthogonal group.*—Let $\mathfrak{G}$ now be the group of all orthogonal transformations with determinants $+1$ or -1. The fundamental theorem of invariant theory for this case states that every invariant J, (19), is a polynomial in the inner products

$$(\mathfrak{x}(i), \mathfrak{x}(k)), \qquad (\mathfrak{x}(i), \mathfrak{u}(k)), \qquad (\mathfrak{u}(i), \mathfrak{u}(k)).$$

Every term of this polynomial must contain each of the vectors $\mathfrak{u}(1), \mathfrak{u}(2), \cdots,$ $\mathfrak{u}(f)$, $\mathfrak{x}(1), \mathfrak{x}(2), \cdots, \mathfrak{x}(f)$ exactly once. Therefore, J is a linear combination of the products of the form,

$$J = (\mathfrak{v}(1), \mathfrak{v}(2))(\mathfrak{v}(3), \mathfrak{v}(4)) \cdots (\mathfrak{v}(2f-1), \mathfrak{v}(2f)), \tag{38}$$

where $\mathfrak{v}(1), \mathfrak{v}(2), \cdots, \mathfrak{v}(2f)$ form a permutation of $\mathfrak{u}(1), \cdots, \mathfrak{u}(f), \mathfrak{x}(1), \cdots,$ $\mathfrak{x}(f)$. We represent $\mathfrak{u}(1), \mathfrak{u}(2), \cdots, \mathfrak{u}(f)$ by f dots in a row, and $\mathfrak{x}(1), \mathfrak{x}(2), \cdots,$ $\mathfrak{x}(f)$ by f dots in a second row. We connect two dots by a line, if the inner product of the corresponding vectors appears in (38). We thus obtain symbols S of the following type (e.g. $f = 5$)

(39) 

To every such symbol S corresponds an invariant (38) which will be denoted by J_S. For instance, the symbol (39) corresponds to

$$(\mathfrak{u}(1), \mathfrak{u}(3))(\mathfrak{u}(2), \mathfrak{x}(1))(\mathfrak{u}(4), \mathfrak{x}(2))(\mathfrak{u}(5), \mathfrak{x}(5))(\mathfrak{x}(3), \mathfrak{x}(4)). \tag{40}$$

The total number of symbols S is

$$N = (2f-1)\cdot(2f-3) \cdots 5\cdot 3\cdot 1, \tag{41}$$

since we have $2f - 1$ possibilities of joining the first dot to another dot, then $2f - 3$ possibilities to join the next unconnected point etc.

Every invariant J, (19), is a linear combination of the N invariants J_S. According to §2, every element B of B is a linear combination of the elements B_S corresponding to J_S. We form B_S.

Put again $\delta(i, k) = 0$ for $i \neq k$, $\delta(i, i) = 1$. The invariant (40) reads, when expressed by the components,

$$J_S = \sum_{i_1, \cdots, i_5, k_1, \cdots, k_5} \delta(i_1, i_3)\delta(i_2, k_1)\delta(i_4, k_2)\delta(i_5, k_5)\delta(k_3, k_4)u^{(i_1)}(1) \cdots u^{(i_5)}(5)x_{k_1}(1) \cdots x_{k_5}(5).$$

In the general case, we associate f indices $i_1, i_2, \cdots, i_f$ with the dots in the upper row, and f indices $k_1, k_2, \cdots, k_f$ with the dots in the lower row, and have

$$J_S = \sum_{i_1, \cdots, i_f, k_1, \cdots k_f} \delta(j_1, j_2)\delta(j_3, j_4) \cdots \delta(j_{2f-1, 2f})u^{(i_1)}(1) \cdots u^{(i_f)}(f)x_{k_1}(1) \cdots x_{k_f}(f),$$

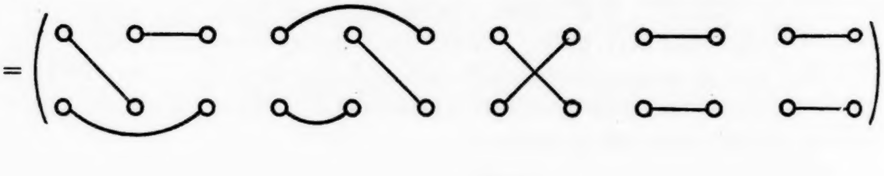
where $j_1, j_2, \cdots j_{2f}$ are the indices $i_1, i_2, \cdots, i_f, k_1, k_2, \cdots, k_f$ arranged in such a way that the dots corresponding to $j_{2\rho-1}, j_{2\rho}$ are connected in S.

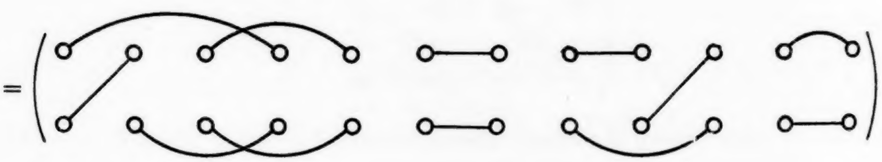
The formulas (19) and (7) then show that B_S is given by

$$B_S = (\delta(j_1, j_2)\delta(j_3, j_4) \cdots \delta(j_{2f-1}, j_{2f})). \tag{42}$$

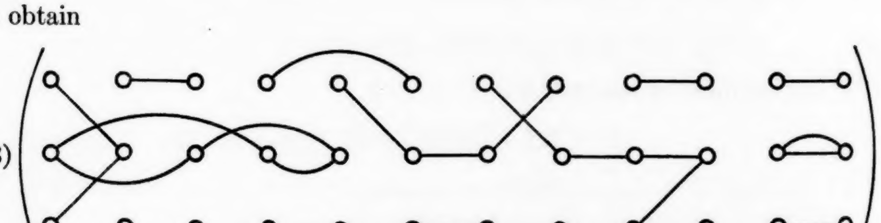
We use the symbol S itself to denote the element B_S. We now have a multiplication of the symbols S.

Using the form (42) of these matrices we obtain by a simple computation the following rule for the multiplication of two of our symbols S_1 and S_2: We move S_2 so that its upper row covers the lower row of S_1. For instance, for $f = 12$ and

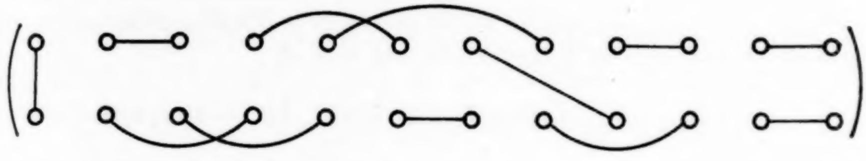
$S_1 =$ 

$S_2 =$ 

we obtain

(43) 

We construct a new symbol S_3 by removing the dots in the middle row in (43) but retaining the dots in the first and last row. Two such dots are to be joined if they are joined in (43). In our example,

$S_3 =$ 

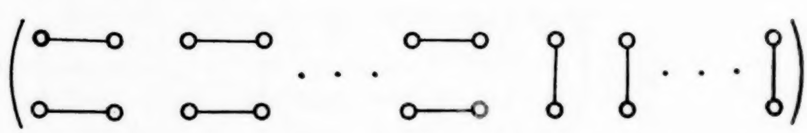
By a cycle in (43) we understand a set of dots $D_1, D_2, \cdots, D_h$ in the middle row in (43) such that D_1 is joined to D_2, D_2 to $D_3, \cdots, D_{h-1}$ to D_h and D_h to D_1. Let k be the total number of such cycles in (43), (in our example $k = 2$). We then have to set

$$S_1 S_2 = n^k S_3. \tag{44}$$

It may happen that the N elements S are linearly dependent in B. We consider the N symbols S as basis elements of a new algebra Γ of order N and define multiplication by (44). Then B is a representation of Γ (but not necessarily a (1-1)-representation). It is easy to show that Γ is associative.

An element S in which dots of the upper row are always connected with dots in the lower row can be interpreted as a permutation of $1, 2, \cdots, f$. The rule for multiplication of the symbols S becomes in this case identical with the rule for multiplication of permutations. The group ring Σ_f of the symmetric permutation group $\mathfrak{S}_f$ in f symbols may, therefore, be considered as a subalgebra of Γ. The particular invariants J_S corresponding to permutations S have the property that only the inner products of a cogredient with a contragredient vector occur. If we had taken the full linear group instead of $\mathfrak{G}$, we would have had to consider only these S. This shows again that the algebra B in case of the full linear group is homomorphic to the group ring Σ_f, (cf. §4).

We now consider the elements S in which there are exactly r connecting lines joining dots of the upper row. The totality of all linear combinations of these S shall be denoted by H_r. In particular, H_0 is identical with Σ_f. If S_1 in (44) belongs to H_r, then S_3 belongs to $\mathrm{H}_{r'}$ with $r' \geqq r$ because every line connecting dots of the upper row in S_1 also appears in S_3. Similarly, if S_2 belongs to H_r, then S_3 belongs to $\mathrm{H}_{r'}$ with $r' \geqq r$. It follows that the sum of $\mathrm{H}_r, \mathrm{H}_{r+1}, \mathrm{H}_{r+2}, \cdots$ forms an ideal. We put

(45) $Z_r =$ 

where there are r horizontal lines in each row. Every symbol S belonging to H_r can be obtained in the form PZ_rQ where P and Q are permutations (elements

S of H_0). Furthermore, we have $Z_r Z_s = n^r Z_s$ for $r \leqq s$. It follows that Z_1 together with the permutations generate the whole algebra Γ. The corresponding fact, of course, is true in the homomorphic algebra B.

The algebra A consists of all the matrices A, commutative with all elements of B. Because of the structure of B it is necessary and sufficient to require that A is commutative with all permutations and with Z_1,

$$Z_1 = (\delta(i_1, i_2)\delta(k_1, k_2)\delta(i_3, k_3)\delta(i_4, k_4) \cdots \delta(i_f, k_f)).$$

The first condition leads to (25); and so we obtain

THEOREM: *A necessary and sufficient condition that* $A = (a([i], [k]))$ *belong to the enveloping algebra of the tensor representation of rank f in case of the orthogonal group is expressed by the equations*

$$a([i]_P, [k]_P) = a([i], [k]) \tag{46 a}$$

or every permutation P, together with

$$\begin{aligned} &\delta(k_1, k_2) \sum_j a((i_1, i_2, \cdots, i_f), (j, j, k_3, k_4, \cdots, k_f)) \\ &\qquad = \delta(i_1, i_2) \sum_j a((j, j, i_3, i_4, \cdots, i_f), (k_1, k_2, \cdots, k_f)) \end{aligned} \tag{46 b}$$

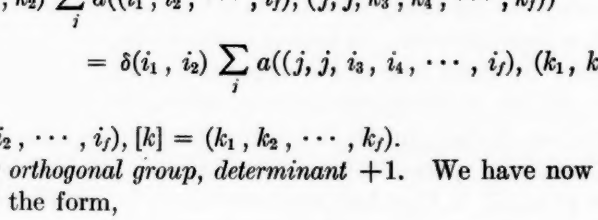
for all $[i] = (i_1, i_2, \cdots, i_f)$, $[k] = (k_1, k_2, \cdots, k_f)$.

b) *The proper orthogonal group, determinant* +1. We have now to consider invariants J_λ of the form,

$$J_\lambda = |\, \mathfrak{v}(1), \mathfrak{v}(2), \cdots, \mathfrak{v}(n) \,|\, (\mathfrak{v}(n+1), \mathfrak{v}(n+2)) \cdots (\mathfrak{v}(2f-1), \mathfrak{v}(2f))$$

besides the invariants (34). Here $|\, \mathfrak{v}(1), \mathfrak{v}(2), \cdots, \mathfrak{v}(n) \,|$ denotes the determinant of the n vectors and $\mathfrak{v}(1), \mathfrak{v}(2), \cdots, \mathfrak{v}(2f)$ form a permutation of $\mathfrak{u}(1), \cdots, \mathfrak{u}(f), \mathfrak{x}(1), \cdots, \mathfrak{x}(f)$. Of course, such an J_λ can be formed only if n is even. In the case of an odd number of dimensions n, the algebras A and B are identical with the corresponding algebras in the case of the full orthogonal group (case a)).

In the case of an even $n = 2\nu$, we have to consider new symbols T besides the symbols S. The symbols T are of the same form as the symbols S, but n dots in them are left unconnected; e.g.

$T_0 =$ 

In the corresponding invariant, the determinant of the n unconnected vectors appears as a factor. We assign a definite order to these vectors, by taking first those in the top row from left to right and then the vectors in the bottom row in the same order. The rule for the multiplication of a symbol T and a symbol S can easily be given. In particular, one proves without difficulty that every symbol T can be represented in the form $S_1 T_0 S_2$ where S_1 and S_2 are symbols S. It follows that Z_1, T_0 together with the permutations generate the whole algebra B in this case. The matrix corresponding to T_0 is

$$(\chi(i_1, i_2, \cdots, i_\nu, k_1, k_2, \cdots, k_\nu)\delta(i_{\nu+1}, k_{\nu+1}) \cdots \delta(i_f, k_f))$$

where $\chi(\rho_1, \rho_2, \cdots, \rho_n)$ is defined in (37).

We therefore obtain for an element A of A:

$$\sum_{j_1,\cdots,j_\nu=1}^{n} a((i_1, \cdots, i_f), (j_1, \cdots, j_\nu, k_{\nu+1}, \cdots, k_f))\chi(j_1, \cdots, j_\nu, k_1, \cdots, k_\nu)$$
$$= \sum_{j_1,\cdots,j_\nu=1}^{n} a((j_1, \cdots, j_\nu, i_{\nu+1}, \cdots, i_f), (k_1, \cdots, k_f))\chi(i_1, \cdots, i_\nu, j_1, \cdots, j_\nu). \tag{46 c}$$

(46 c) together with (46 a) and (46 b) characterizes the matrices of A in this case.

The rule for the multiplication of the two elements T can also be formulated. It is, however, more complicated and shall not be given here.

c) *The case of the complex-group.* As in §3, we denote the coordinates by $x_1, x_{1'}, x_2, x_{2'}, \cdots, x_\nu, x_{\nu'}$; $n = 2\nu$, and take again the invariant bilinear form $\{\mathfrak{x}, \mathfrak{y}\}$ in the form (24). The basis invariants from which we have to compose the invariants J_κ are here

$$\{\mathfrak{x}(i), \mathfrak{x}(k)\}, \qquad (\mathfrak{u}(j), \mathfrak{x}(l)), \qquad \{\mathfrak{u}(i), \mathfrak{u}(k)\}, \qquad (i < k).$$

Again we may denote the J_S by means of the symbols S. But in cases of two connected dots lying in the same row, we have to consider the skew product $\{\mathfrak{u}(i), \mathfrak{u}(k)\}$, respectively $\{\mathfrak{x}(i), \mathfrak{x}(k)\}$, instead of the inner product. So the invariant J_S corresponding to S, (39), is here (instead of (40))

$$J_S = \{\mathfrak{u}(1), \mathfrak{u}(3)\}(\mathfrak{u}(2), \mathfrak{x}(1))(\mathfrak{u}(4), \mathfrak{x}(2))(\mathfrak{u}(5), \mathfrak{x}(5))\{\mathfrak{x}(3), \mathfrak{x}(4)\}.$$

The corresponding matrix is

$$B_S = (\epsilon(i_1, i_3)\delta(i_2, k_1)\delta(i_4, k_2)\delta(i_5, k_5)\, \epsilon(k_3, k_4)).$$

In case of two arguments i or two arguments k, the factor δ in (42) has to be replaced by ϵ. The indices i_ρ, k_ρ range over $1, 1', 2, 2', \cdots, \nu, \nu'$. The law of multiplication is similar to (44). Again, S_3 has to be formed in the same way as in the case of the orthogonal group. One has, however, to add a factor

$\varphi(S_1, S_2)$ on the right side, whose value is $+1$, -1 or 0. Again Z_1, (45), together with the permutations generates B. We then obtain

$$(47)\quad \begin{cases} a([i]_P, [k]_P) = a([i], [k]), \\ \epsilon(k_1, k_2) \sum_{j_1, j_2 = 1, 1', \cdots, \nu, \nu'} a((i_1, \cdots, i_f), (j_1, j_2, k_3, \cdots, k_f))\epsilon(j_1, j_2) \\ = \epsilon(i_1, i_2) \sum_{j_1, j_2 = 1, 1', \cdots, \nu, \nu'} a((j_1, j_2, i_3, \cdots, i_f), (k_1, \cdots, k_f))\epsilon(j_1, j_2). \end{cases}$$

as necessary and sufficient conditions that $A = (a([i], [k]))$ belongs to A. This result is equivalent to the one given by Weyl, but it does not involve the algebras A belonging to tensors of other ranks than f.

THE UNIVERSITY OF TORONTO.

Annals of Mathematics
Vol. 38, No. 4, October, 1937

, (45),

).

This
gebras

GEOMETRY OF CONFORMAL SYMMETRY (SCHWARZIAN REFLECTION)[1]

By Edward Kasner

(Received April 6, 1937)

In this paper we derive certain new geometric theorems connected with the Schwarzian reflection or *conformal symmetry* determined by an arbitrary analytic curve C in the plane. These formulas describe the effect of the symmetry on curvature and some of its arc-length derivatives, for curves crossing the base curve C (theorem I). They may be considered as first steps towards developing a purely geometric construction for the conformal symmetry determined by a general base curve. In the case of a curve tangent to C there is an exceedingly simple *arithmetic mean relationship for curvature* and its first two derivatives (corollary 1). When C is a circle, the symmetry is of course ordinary inversion, and then the set of relations again becomes simpler (corollary 2). An application to the general bisection problem of curvilinear angles is discussed. Finally (§4) we show how the expression for the conformal invariant of a general curvilinear right angle can be deduced from the invariant of a horn angle of second order contact in view of the symmetry formulas. A relation to Mullins' inversive invariant is indicated.

I wish to thank Annette Vassell and George Comenetz for valuable assistance in preparing this paper. The results have been extended to conformal symmetry on any curved surface by Comenetz,[2] who has also found the appropriate conformal invariants.

1. The conformal symmetry in a regular analytic arc C may be defined as the unique anticonformal transformation of a neighborhood of C into itself which leaves every point of C fixed. That such a transformation exists follows from the fact that it is possible to map a neighborhood of a straight line segment L conformally onto a neighborhood of C, with L going into C. The symmetry in C is then simply the transform of the ordinary reflection in L. Any symmetry is involutoric, since its square is a conformal transformation leaving the points of C fixed and hence is the identity. The symmetry in C is unique (in any given neighborhood of C), for the product of two such symmetries must as before be the identity. Clearly the symmetry which a curve determines is a covariant of the curve under any conformal transformation.[3]

---

[1] Abstract in *Science*, **83** (1936), p. 480. The author studied products of symmetries in a paper, *Infinite Groups Generated by Conformal Transformations of Period Two* (*Involutions and Symmetries*), Amer. Jour. Math., **38** (1916), pp. 177–184.

[2] Abstract in Bull. Amer. Math. Soc., **42** (1936), p. 806.

[3] H. A. Schwarz, *Werke*, II, p. 151; Osgood, *Funktionentheorie* (5th ed.), I, p. 706.

Let O be any point of C, and let the coordinate axes be chosen so that O is the origin and the tangent to C at O is the x-axis. Then C will be described in some neighborhood of O by an equation

$$y = s_2x^2 + s_3x^3 + s_4x^4 + \cdots . \tag{1}$$

We shall suppose that C is oriented, say, to the right.

An anticonformal transformation leaving O fixed can be represented in some neighborhood of O by

$$w = m_1\bar{z} + m_2\bar{z}^2 + m_3\bar{z}^3 + \cdots , \tag{2}$$

where $\bar{z} = x - iy$ and $w = X + iY$ (X, Y being the transformed coordinates), and where the coefficients m_n are complex numbers: $m_n = \mu_n + i\nu_n$. If (2) represents the symmetry in C, then the m_n are determined by the coefficients s_k in (1). We shall obtain the expressions for $m_1, \cdots, m_5$ explicitly. To do this we separate (2) into real and imaginary parts:

$$\begin{aligned} X &= \mu_1x + \nu_1y + \mu_2(x^2 - y^2) + 2\nu_2xy + \cdots , \\ Y &= \nu_1x - \mu_1y + \nu_2(x^2 - y^2) - 2\mu_2xy + \cdots . \end{aligned} \tag{2'}$$

Since the points of C are left fixed, equations (2′) become identities in x if we replace X, Y by x, y and then eliminate y by means of (1). Equating coefficients of $x, \cdots, x^5$ respectively and solving for $\mu_1, \nu_1, \cdots, \mu_5, \nu_5$, we find that

$$\begin{aligned} &\mu_1 = 1,\ \nu_1 = 0;\ \mu_2 = 0,\ \nu_2 = 2s_2;\ \mu_3 = -4s_2^2,\ \nu_3 = 2s_3; \\ &\mu_4 = -10s_2s_3,\ \nu_4 = 2s_4 - 10s_2^3;\ \mu_5 = -12s_2s_4 - 6s_3^2 + 28s_2^4, \\ &\nu_5 = 2s_5 - 42s_2^2s_3. \end{aligned} \tag{3}$$
[4]

Let γ be the curvature of C at O, γ' the derivative of curvature with respect to arc-length at O, γ'' the second derivative, and so on. Then the coefficients $s_2, \cdots, s_5$ can be expressed in terms of the metric invariants $\gamma, \cdots, \gamma'''$ of C by the formulas

$$2s_2 = \gamma, \qquad 6s_3 = \gamma', \qquad 24s_4 = \gamma'' + 3\gamma^3, \qquad 120s_5 = \gamma''' + 19\gamma^2\gamma'. \tag{4}$$

Substituting (4) into (3), we obtain $\mu_1, \cdots, \nu_5$ in terms of $\gamma, \cdots, \gamma'''$.

Fundamental Theorem. *The symmetry determined by any curve C is represented explicitly to the fifth order by*

$$\begin{aligned} w = \bar{z} &+ i\gamma\bar{z}^2 + (-\gamma^2 + \tfrac{1}{3}i\gamma')\bar{z}^3 + [-\tfrac{5}{6}\gamma\gamma' + i(\tfrac{1}{12}\gamma'' - \gamma^3)]\bar{z}^4 \\ &+ [(-\tfrac{1}{4}\gamma\gamma'' - \tfrac{1}{6}(\gamma')^2 + \gamma^4) + i(\tfrac{1}{60}\gamma''' - \tfrac{43}{30}\gamma^2\gamma')]\bar{z}^5 + \cdots . \end{aligned} \tag{5}$$

This is the fundamental intrinsic representation of general conformal symmetry.

---

[4] Cf. G. A. Pfeiffer, *On the Conformal Geometry of Analytic Arcs*, Amer. Jour. Math., 37 (1915), p. 399 (with $a_1 = 0$).

2. Let D be any oriented analytic arc intersecting C at O, and let θ be the angle from C to D at O. Then D has an equation

$$y = a_1 x + a_2 x^2 + a_3 x^3 + \cdots, \tag{6}$$

where $a_1 = \tan \theta$.[5] If $\bar{D}$ denotes the image of D under the symmetry in C, then $\bar{D}$ is given by

$$Y = A_1 X + A_2 X^2 + A_3 X^3 + \cdots. \tag{7}$$

We wish to determine the transformed coefficients $A_1, \cdots, A_4$ in terms of the a_n and the coefficients of the symmetry (5).

As before, we convert (5) into the real form (2′). Then we substitute the expressions for X and Y into (7), and finally eliminate y by using (6). In the resulting identity we equate coefficients for the first four powers of x. Solving for $A_1, \cdots, A_4$, we find

$$A_1 = -a_1, \tag{8_1}$$

$$A_2 = -a_2 + \gamma(1 + a_1^2), \tag{8_2}$$

$$A_3 = -a_3 + 4\gamma a_1 a_2 + \tfrac{1}{3}\gamma'(1 - a_1^4) - 2\gamma^2 a_1(1 + a_1^2), \tag{8_3}$$

$$\begin{aligned} A_4 = {} & -a_4 + 6\gamma a_1 a_3 + 3\gamma a_2^2 + \tfrac{1}{3}\gamma' a_1 a_2(3 - 5a_1^2) - 3\gamma^2 a_2(1 + 5a_1^2) \\ & + \tfrac{1}{12}\gamma''(1 - 3a_1^2)(1 + a_1^2) + \tfrac{1}{6}\gamma\gamma' a_1(-9 + 11a_1^2)(1 + a_1^2) + \gamma^3(1 + 5a_1^2)(1 + a_1^2). \end{aligned} \tag{8_4}$$

The first equation shows that the angle from C to $\bar{D}$ is $-\theta$, as indeed it must be since the transformation is anticonformal.

The calculation of A_5 for a general value of a_1 would be too long. However, we will consider two important special cases, namely, when D is tangent to C (*horn angle*) and when D is perpendicular to C (*right angle*).

If D is tangent to C at O, we set $a_1 = 0$ in (6) (and in $(8_1) - (8_4)$), and proceed as before. The result is

$$A_5 = -a_5 + 4\gamma a_3(2a_2 - \gamma) + \gamma' a_2^2 - \tfrac{7}{3}\gamma\gamma' a_2 + \tfrac{1}{60}\gamma''' + \tfrac{37}{30}\gamma^2\gamma' \qquad (\text{when } a_1 = 0). \tag{$8_5'$}$$

If D is perpendicular to C at O, it has an equation of the form

$$x = b_2 y^2 + b_3 y^3 + b_4 y^4 + \cdots, \tag{9}$$

and the equation of $\bar{D}$ is

$$X = B_2 Y^2 + B_3 Y^3 + B_4 Y^4 + \cdots. \tag{10}$$

[5] The calculation which follows does not apply if $\theta = 90°$ or $270°$; nevertheless equations $(11_1) - (11_3)$ are still correct then. For if we had represented D by $x = b_1 y + b_2 y^2 + \cdots$ instead of by (6) we would necessarily have found the same equations.

In this case the usual steps lead to the relation

$$(8_5'') \quad B_5 = -b_5 + 2\gamma(2B_4 - b_4) - 3\gamma^2(2B_3 + b_3) - 2\gamma b_2^2 B_2 + b_2\gamma'(2B_2 - b_2) + \tfrac{1}{6}\gamma''(2b_2 - B_2) + 4\gamma^3(B_2 - b_2) - \tfrac{1}{60}\gamma''' + \tfrac{43}{30}\gamma^2\gamma'.$$

Let Γ denote the curvature of the curve D at O, Γ' the derivative of curvature with respect to arc-length at O, Γ'' the second derivative, and so on; and let $\bar{\Gamma}$, $\bar{\Gamma}'$, $\bar{\Gamma}''$, $\cdots$ be the corresponding quantities at O for the reflected curve $\bar{D}$. Then (assuming first that D is oriented to the right) the coefficients $a_2, \cdots, a_5$ in the equation (6) for D are expressible in terms of $\Gamma, \cdots, \Gamma'''$ and a_1 as follows:

$$\begin{aligned} 2a_2 &= (1 + a_1^2)^{3/2}\Gamma, \\ 6a_3 &= (1 + a_1^2)^2(\Gamma' + 3a_1\Gamma^2), \\ 24a_4 &= (1 + a_1^2)^{5/2}[\Gamma'' + 10a_1\Gamma\Gamma' + 3(1 + 5a_1^2)\Gamma^3], \\ 120a_5 &= \Gamma''' + 19\Gamma^2\Gamma' \quad (\text{when } a_1 = 0). \end{aligned}$$

The corresponding formulas for $\bar{D}$ have $A_1, \cdots, A_5$, $\bar{\Gamma}, \cdots, \bar{\Gamma}'''$ in place of $a_1, \cdots, \Gamma'''$ respectively. If D is perpendicular to C and oriented upwards, the appropriate formulas are

$$2b_2 = -\Gamma, \qquad 6b_3 = -\Gamma', \qquad 24b_4 = -\Gamma'' - 3\Gamma^3, \qquad 120b_5 = -\Gamma''' - 19\Gamma^2\Gamma'.$$

For $\bar{D}$ (oriented downwards) we replace $b_2, \cdots, b_5, \Gamma, \Gamma', \Gamma'', \Gamma'''$ by $B_2, \cdots, B_5$, $-\bar{\Gamma}, +\bar{\Gamma}', -\bar{\Gamma}'', +\bar{\Gamma}'''$ respectively.

Now using the above relations and the fact that $a_1 = \tan\theta$, we eliminate the a_n, A_n, b_n and B_n from $(8_2) - (8_4)$, $(8_5')$ and $(8_5'')$ (for $(8_5'')$ we use $(11_1) - (11_3)$ with $\theta = 90°$). In the end we find the equations

$$(11_1) \qquad \bar{\Gamma} + \Gamma = 2\gamma\cos\theta,$$

$$(11_2) \qquad \bar{\Gamma}' + \Gamma' = 2\gamma'\cos 2\theta,$$

$$(11_3) \qquad \bar{\Gamma}'' + \Gamma'' = 2\gamma''\cos 3\theta - 4\Gamma\gamma'\sin 2\theta + 4\gamma(\gamma' + \Gamma')\sin\theta,$$

which hold for any value of θ.[6]

For the special value $\theta = 0°$ we have the further equation

$$(11_4') \qquad \bar{\Gamma}''' + \Gamma''' = 2\gamma''' - 4\gamma'(\gamma - \Gamma)(\gamma - 2\Gamma) - 4\,\Gamma'\gamma(\gamma - \Gamma)$$

and, for $\theta = 90°$,

$$(11_4'') \quad \bar{\Gamma}''' + \Gamma''' = 2\gamma''' + 10(\Gamma\gamma'' + \Gamma''\gamma) - 4(3\gamma^2\gamma' - 2\Gamma^2\gamma' + 5\Gamma'\gamma^2).$$

To summarize our main result we state the following theorem.

---

[6] If the orientation of a curve is reversed, the curvature and its even derivatives change sign, but the odd derivatives are unchanged. It follows easily that $(11_1) - (11_3)$ are still correct if D is oriented to the left. If we replace γ, γ'' by $-\gamma, -\gamma''$, $(11')$ applies for $\theta = 180°$ and $(11_4'')$ for $\theta = 270°$.

THEOREM I. *Let C be a given analytic curve, let O be a point of it, and let D be any curve cutting C at O at an angle θ. Let the image of D under the conformal symmetry determined by C be the curve $\bar{D}$; then $\bar{D}$ cuts C at O at the angle $-\theta$. Denote the values at O of curvature and its successive arc-length derivatives, for C by $\gamma, \gamma', \cdots$, for D by $\Gamma, \Gamma', \cdots$, and for $\bar{D}$ by $\bar{\Gamma}, \bar{\Gamma}', \cdots$. Then these curvatures and their derivatives up to the second order are related, for any value of θ, by the set of fundamental formulas $(11_1) - (11_3)$. In addition we have for the third derivatives the formulas $(11_4')$ and $(11_4'')$, which apply in the special cases $\theta = 0°$ and $\theta = 90°$ respectively.*[7]

The remarkably simple forms of (11_1) and (11_2) suggest simple geometric statements *analogous to Meusnier's theorem.* They, and the other formulas, also supply geometric information concerning the curvature and derivatives of curvature of the *curve bisecting a general curvilinear angle,* as discussed later.

3. The formulas are particularly simple when $\theta = 0°$. This may be stated separately:

COROLLARY 1: *Suppose that a curve D is tangent to (and similarly sensed with) the base curve of a conformal symmetry. Then at the point of tangency the curvature of the base curve equals the arithmetic mean of the curvatures of D and the reflection of D under the symmetry. The same relation holds for the first derivatives of curvature with respect to arc-length, and for the second derivatives; but for the third there is the more complicated relation $(11_4')$.*

If the base curve C is a circle we have the important special case of *inversion* (transformation by reciprocal radii vectores). To adapt the formulas (11) to this case we need only make $\gamma = \pm 1/r$ (r = radius of circle) and $\gamma' = \gamma'' = \gamma''' = 0$.

COROLLARY 2: *Under inversion in a circle, curvature and its derivatives for any curve at a point of crossing the circle obey the relations*

$$\bar{\Gamma} + \Gamma = 2\gamma \cos \theta,$$

$$\bar{\Gamma}' + \Gamma' = 0,$$

$$\bar{\Gamma}'' + \Gamma'' = 4\Gamma'\gamma \sin \theta,$$

which are valid for any θ; and also

$$\bar{\Gamma}''' + \Gamma''' = 4\Gamma'\gamma(\Gamma - \gamma) \textit{ when } \theta = 0°, \textit{ and}$$

$$\bar{\Gamma}''' + \Gamma''' = 10\gamma(\Gamma'' - 2\Gamma'\gamma) \textit{ when } \theta = 90°,$$

where γ is the curvature of the circle.[8]

---

[7] We proved this for curves D analytic at O, but then it must also be true if D merely possesses a fifth derivative at O. This is seen by drawing an analytic curve having fifth order contact with the differentiable curve.

[8] If we had calculated the $\bar{\Gamma}''' + \Gamma'''$ formula for a general θ, it would be easy to show, by

Returning to the case of a general base curve C, let us think of D and $\bar{D}$ as given, while C is to be found. This is the *bisection problem*: to draw a curve C through the vertex of a given curvilinear angle in such a way that the sides of the angle are reflections of each other under the symmetry in C.[9]

Let the magnitude of the angle from $\bar{D}$ to D be α; then formulas (11) apply at once, with $\theta = \frac{1}{2}\alpha$ (or $\frac{1}{2}\alpha + \pi$, depending on the orientation of the bisector C). To determine the bisector we simply solve for $\gamma, \gamma', \cdots$, if it is possible to do so.

It can be shown that the general relation of the type $(11_1) - (11_3)$ has the form

$$\bar{\Gamma}^{(n)} + \Gamma^{(n)} = 2\gamma^{(n)} \cos{(n+1)\theta} + \cdots,$$

$(n = 0, 1, 2, \cdots; \gamma^{(0)} \equiv \gamma$, etc.), where the remaining terms on the right involve $\gamma, \cdots, \gamma^{(n-1)}, \Gamma, \cdots, \Gamma^{(n-1)}$ and θ. Hence if $\cos{(n+1)\theta}$ never vanishes $(n = 0, 1, \cdots)$, all of the quantities $\gamma, \gamma', \cdots$ can be determined, and a unique (formal) bisector exists.[10] That is, *if* $\alpha = (p/q)\pi$ *where* p *is even* (p/q being in lowest terms), *or if* α/π *is irrational, there is a unique conformal bisector. Otherwise there is in general no bisector.* For instance, angles of magnitude $0, 2\pi/3, 2\pi/5, 4\pi/5, \cdots$ have a unique bisector, while angles of magnitude $\pi, \pi/2, \pi/3, \pi/4, 3\pi/4, \pi/5, \cdots$ have in general no bisector. Thus if $\alpha = (p/q)\pi$ where q is even $(\pi/2, \pi/4, 3\pi/4, \cdots)$, neither an internal nor an external bisector exists in general; while if q is odd $(0, \pi, \pi/3, \cdots)$ there is a bisector, in general unique, internal if p is even and external if p is odd.[11]

4. Let C_1, C_2 be a pair of oriented analytic curves passing through a point O, such that the angle from C_1 to C_2 at O is 90°. Let $\gamma_1, \gamma_1', \cdots$ and $\gamma_2, \gamma_2', \cdots$ denote curvature and its arc-length derivatives at O for C_1 and C_2 respectively. Suppose that $\gamma_2' + \gamma_1' \neq 0$; then by definition C_1, C_2 form a *general curvilinear right angle*.[12]

We reflect C_1 by conformal symmetry in C_2, obtaining an image curve $\bar{C}_1$.

---

eliminating γ and θ and allowing for magnification and reflection, that the *"inversive curvature"*

$$(4\Gamma'\Gamma''' - 5\Gamma''^2 - 4\Gamma^2\Gamma'^2)/\Gamma'^3$$

is the invariant of lowest order of a single curve under the (direct) inversion group. This fundamental inversive invariant of a curve was first obtained by G. W. Mullins in a different form by using Lie theory. Mullins also finds an interesting geometric interpretation. See *Differential Invariants under the Inversion Group*, Columbia University Dissertation, 1917; also B. C. Patterson, *The Differential Invariants of Inversive Geometry*, Amer. Jour. Math., **50** (1928), pp. 553–568.

The first equation in corollary 2 is essentially equivalent to an elementary formula for the effect of inversion on the radius of a circle which cuts the base circle at an angle θ.

[9] E. Kasner, *Conformal Geometry*, Proc. Fifth Inter. Congr. Math., Cambridge, 1912, II, p. 81.

[10] "Formal," because the series for C may not converge; see Pfeiffer, *The Functional Equation* $f[f(x)] = g(x)$, Annals of Math., **20** (1918), p. 13.

[11] Cf. Pfeiffer, *On the Conformal* $\cdots$, p. 419, for further results.

[12] The reference for this section is: E. Kasner, *The Two Conformal Invariants of Fifth Order*, read at the Inter. Congr. Math., Oslo, 1936, to be published in *Trans. Amer. Math. Soc.*, 1937.

Let $\bar{C}_1$ with its orientation reversed be called C_3. Denote curvature and its derivatives at O by $\bar{\gamma}_1, \bar{\gamma}_1', \cdots$ for $\bar{C}_1$, and by $\gamma_3, \gamma_3', \cdots$ for C_3. Then we have

$$\gamma_3 = -\bar{\gamma}_1, \qquad \gamma_3' = \bar{\gamma}_1', \qquad \gamma_3'' = -\bar{\gamma}_1'', \qquad \gamma_3''' = \bar{\gamma}_1'''. \tag{12}$$

Furthermore, since the angle from C_2 to C_1 is 270°, formulas $(11_1) - (11_3)$ and $(11_4'')$ (as modified in the footnote) yield

$$\begin{gathered} \bar{\gamma}_1 = -\gamma_1, \qquad \bar{\gamma}_1' = -\gamma_1' - 2\gamma_2', \qquad \bar{\gamma}_1'' = -\gamma_1'' - 4\gamma_2(\gamma_2' + \gamma_1'), \\ \bar{\gamma}_1''' = -\gamma_1''' + 2\gamma_2''' - 10(\gamma_1\gamma_2'' + \gamma_1''\gamma_2) - 4(3\gamma_2^2\gamma_2' - 2\gamma_1^2\gamma_2' + 5\gamma_1'\gamma_2^2). \end{gathered} \tag{13}$$

Now C_1 and C_3 are tangent at O, with like orientation; and we see that $\gamma_3 = \gamma_1$, and $\gamma_3' - \gamma_1' = -2(\gamma_2' + \gamma_1') \neq 0$. Consequently C_1, C_3 form what we have called a *horn angle of second order contact*, or an "H_3".

We proved in the reference that under any (direct) conformal transformation the value of[13]

$$\frac{4(\gamma_3' - \gamma_1')(\gamma_3''' - \gamma_1''') - 5(\gamma_3'' - \gamma_1'')^2 - 4\gamma^2(\gamma_3' - \gamma_1')^2}{(\gamma_3' - \gamma_1')^3} \tag{14}$$

$(\gamma \equiv \gamma_3 = \gamma_1)$, and the sign of $\gamma_3' - \gamma_1'$, are *invariants of an* H_3. If we eliminate $\gamma_3, \gamma_3', \cdots$ by means of (12), and then $\bar{\gamma}_1, \bar{\gamma}_1', \cdots$ by means of (13), we obtain the expression

$$\frac{(\gamma_2''' - \gamma_1''') - 5(\gamma_1\gamma_2'' + \gamma_1''\gamma_2) - \gamma_2^2\gamma_2' + 5\gamma_1^2\gamma_2' - 5\gamma_1'\gamma_2^2 + \gamma_1^2\gamma_1'}{(\gamma_2' + \gamma_1')^2} \tag{15}$$

(times 2). Since the relation of symmetry is preserved by a conformal transformation, we conclude that the *value of the above expression* and the sign of $\gamma_2' + \gamma_1' (= -\frac{1}{2}(\gamma_3' - \gamma_1'))$ are *conformal invariants of a general curvilinear right angle.*

We remark, in conclusion, that (14) and (15) are the only *absolute* conformal invariants of fifth order which arise in conformal differential geometry. There are no invariants of fourth order, and the only invariant of third order is the one (stated in the preceding footnote) for an ordinary horn angle, namely,

$$I_{12} = \frac{\gamma_2' - \gamma_1'}{(\gamma_2 - \gamma_1)^2}.$$

The denominator of this fraction, or rather the equivalent expression $|\gamma_2 - \gamma_1|$, was rediscovered as a relative conformal invariant by Ostrowski (Jahr. Deut. Math.-Ver., 1934) who gives many interesting applications. The numerators and denominators of our absolute invariants are all *relative invariants.* There exist two relative invariants of fourth order, in addition to the relative invariants of second, third, and fifth orders given in the present paper.

COLUMBIA UNIVERSITY.

[13] For a horn angle H_2 of *first order contact* the fundamental invariant is of third order: $I_{12} = \dfrac{\gamma_2' - \gamma_1'}{(\gamma_2 - \gamma_1)^2}$; and the reciprocal M_{12} is defined as the *natural measure* of the horn angle. See *Cambridge Congr.*, 1912 and *Proc. Nat. Acad.*, **22** (1936) p. 303, **23** (1937) p. 337, and *Science*, **85** (1937), p. 480. The last two papers deal with *Trihornometry*.

Annals of Mathematics
Vol. 38, No. 4, October, 1937

CLOSURES AND ADJOINTS OF LINEAR DIFFERENTIAL OPERATORS

By Israel Halperin

(Received July 2, 1936)

Introduction

In a real or complex Banach space B of elements $f, g, \cdots$, with norm $\| f \|, \| g \|, \cdots$, an *operator* T is a function whose values lie in the same Banach space but which need not be everywhere defined nor single-valued.[1] T_2 is said to be an *extension* of T_1, in symbols $T_2 \supset T_1$, or $T_1 \subset T_2$, if for every f, the set of values of $T_2 f$ includes all values of $T_1 f$. The operator T is called *linear* if the set of values of $T(cf + dg)$ includes all values of $cTf + dTg$, for all f, g and all numbers c, d (real or complex, depending on B). A *closure* operator $\tilde{T}$ (for arbitrary T) is defined by the condition that $\tilde{T}f$ has g as one of its values if and only if

$$f_m \to f, \qquad Tf_m \to g, \qquad \text{as } m \to \infty,$$

hold simultaneously for some sequence of elements f_m and values Tf_m. It follows at once that $T_2 \supset T_1$ implies $\tilde{T}_2 \supset \tilde{T}_1$ and that always $\tilde{T} \supset T$. If $\tilde{T} = T$ the operator T is called *closed*. It can be shown that every $\tilde{T}$ is closed, that $\tilde{T}$ is linear if T is linear, and that single-valuedness of T is not a sufficient

---

[1] See S. Banach: *Opérations Linéaires*. The theory given there for real B spaces extends in an obvious way to complex spaces. The set of f's for which Tf is a non-vacuous set is called the *domain* of T. A sub-set M of a real (resp. complex) B space, is called *linear* if f, g lie in M implies $f + g$ and cf lie in M for all real (resp. complex) c.

[2] Let the direct sum space $B \oplus B$ of all couples $< f_1, f_2 >$, f_1, f_2 in B, be made into a Banach space under the definitions

$$< f_1, f_2 > + < g_1, g_2 > = < f_1 + g_1, f_2 + g_2 >$$
$$c < f_1, f_2 > = < cf_1, cf_2 >$$
$$\| < f_1, f_2 > \| = (\| f_1 \|^2 + \| f_2 \|^2)^{\frac{1}{2}}.$$

If T is any operator in B, then the set of all $< f, Tf >$ defines a sub-set $G(T)$ in $B \oplus B$ called the *graph* of T. Every sub-set of $B \oplus B$ is obviously the graph of a uniquely defined operator, so that there is a (1,1) correspondence between the operators in B and the sub-sets of $B \oplus B$. It is easy to see that T is linear if and only if $G(T)$ is linear, and that $G(\tilde{T})$ is the topological closure of $G(T)$. Hence $\tilde{T}$ is always closed and $\tilde{T}$ is linear if T is so.

For one-dimensional real Banach space, the real numbers, the linear, single-valued T's with non-single-valued $\tilde{T}$, are precisely the well-known non-Lebesgue-measurable solutions of $f(x + y) = f(x) + f(y)$ and can be determined by means of a Hamel's basis for the real numbers, (see G. Hamel: Mathematische Annalen, vol. 60 (1905), pp. 459–462). A similar Hamel's basis can be constructed for any Banach space and used to determine the everywhere defined, linear and single-valued T's with non-single-valued closures.

condition for that of $\tilde{T}$.[2] For linear, single-valued and everywhere defined T, it is known that, T is closed is equivalent to T is continuous.[3]

If the Banach space is restricted to be a Hilbert space H with inner product (f, g) and norm $\|f\| = (f, f)^{\frac{1}{2}}$, then an *adjoint* operator T^* is defined as follows: T^*g has an element g^* as one of its values if and only if

$$(Tf, g) = (f, g^*)$$

for all f, Tf.[4] From this definition follows that $T_2 \supset T_1$ implies $T_1^* \supset T_2^*$. It can also be shown that T^* is always linear and closed, that $(\tilde{T})^* = T^*$ for all T, that $T^{**} \equiv (T^*)^* = \tilde{T}$ if T is linear, and that T^* is single-valued if and only if the 0 element in H is the only element orthogonal to all elements in the domain of T.[5]

In this paper we shall be concerned with a linear differential expression

$$\sum_{r_1+\cdots+r_t \leq n} p_{r_1,\cdots,r_t}(x_1, \cdots, x_t) \frac{\partial^{r_1+\cdots+r_t}}{\partial x_1^{r_1} \cdots \partial x_t^{r_t}}$$

considered as determining an operator in a Hilbert space of functions which are Lebesgue square summable (abbrev. L.s.s.) with inner product $(f, g) = \int f(P)\overline{g(P)}\, dP$. We shall seek to determine explicitly the closure and adjoint operators. Special cases of this problem are treated in Chapter III of Stone's

---

[3] See Banach loc. cit. page 41, Theorem 7. Everywhere defined continous operators are necessarily closed, but the converse may be false for non-linear ones, as shown by the function $f(x) = 1/x$ for $x \neq 0$, $f(0) = 0$.

[4] For the theory of Hilbert space, see J. v. Neumann: Mathematische Annalen 102 (1929) pp. 49–131, or M. H. Stone: *Linear Transformations in Hilbert Space*, American Mathematical Society Colloquium Publications, vol. XV, 1932. For a discussion of adjoint and closure operators see particularly J. v. Neumann: *Lecture Notes on Operator Theory*, mimeographed at Princeton, N. J., 1934.

f is said to be *orthogonal* to g if $(f, g) = 0$. If M is any sub-set of H, let $\{M\}$ be the set of all finite sums $\sum c_i f_i$, f_i in M, c_i arbitrary numbers. Then $\{M\}$ is the smallest linear sub-set of H containing M, and the topological closure of $\{M\} \equiv [M]$ say, is the smallest closed linear space containing M. (If M is finite dimensional $[M] = \{M\}$.) Let $\ominus M$ be the set of those f which are orthogonal to all elements in M. Then $\ominus M$ is always linear and closed, and $\ominus\ominus M = [M]$, $\ominus M = \ominus[M]$. An everywhere defined linear operator U whose values cover all of H, is called unitary if $(Uf, Ug) = (f, g)$ for all f, g. For unitary U, $\ominus U(M) = U(\ominus M)$.

[5] Let $H \oplus H$ (see (2)) be made into a Hilbert space with inner product $(\langle f_1, f_2 \rangle, \langle g_1, g_2 \rangle) = (f_1, g_1) + (f_2, g_2)$. Then $T^*g = g^*$ means $(Tf, g) = (f, g^*)$, that is, $(\langle Tf, -f \rangle, \langle g, g^* \rangle) = 0$ for all f, Tf. Define a unitary operator U in $H \oplus H$ by $U\langle f_1, f_2 \rangle = \langle f_2, -f_1 \rangle$. Then $G(T^*) = \ominus U(G(T))$. Hence T^* is linear and closed. $G(T^*) = \ominus U(G(T)) = U(\ominus G(T)) = U(\ominus[G(T)]) = U(\ominus G(\tilde{T})) = \ominus U(G(\tilde{T})) = G((\tilde{T})^*)$ gives $T^* = \tilde{T}^*$. If $G(T)$ is linear, $UU(G) = G$ and so $G(T^{**}) = \ominus U(\ominus U(G(T))) = \ominus\ominus UU(G(T)) = [G(T)] = G(\tilde{T})$, that is, if T is linear then $T^{**} = \tilde{T}$. Finally, $(Tf, g) = (f, g^*) = (f, g^{**})$ implies $(f, g^* - g^{**}) = 0$. If now $N = \ominus$ (domain of T) consists of only the 0 element, then $g^* - g^{**} = 0$, and hence T^* is single-valued. But if $h \neq 0$ is in N, then $T^*(0)$ clearly has 0 and h as values, and T^* is not single-valued.

treatise.[4] Other writers have been able to avoid the real variable difficulties of determining the closure and adjoint by using a different inner product,[6] but it is desirable to consider the operator in classical Hilbert space, which will be done here. Since the differential expression can be applied directly only to functions of sufficiently high order of differentiability, a linear space of such functions, extensive enough to be dense in the Hilbert space, will be chosen as domain of the operator and then the closure and adjoint will be sought.

In §1 we consider the ordinary differential operational expression

$$p_n(x) \frac{d^n \cdot}{dx^n} + \cdots + p_0(x)$$

with linear boundary conditions, on a finite interval (a, b). It is first assumed that $|p_n(x)| \geqq \epsilon > 0$ and that the $p_r(x)$ satisfy certain differentiability conditions, and the operator is taken as defined for polynomials on (a, b). The closure is then shown to be single-valued and defined for exactly those functions $f(x)$ which satisfy the given boundary conditions and have absolutely continuous (abbrev. abs. cont.) $(n-1)^{\text{th}}$ derivative functions and L.s.s. n^{th} derivative functions; the value of the closure at $f(x)$ is

$$\sum_{r=0}^{n} p_r(x) \frac{d^r f(x)}{dx^r},$$

which is defined almost everywhere and is L.s.s. The closure can therefore be represented directly by the original differential expression. The adjoint turns out to be a differential operator associated in the same way with the so-called Lagrange-Adjoint system.[7] If the $p_r(x)$ are required to be merely measurable and bounded, then the adjoint need no longer be again a differential operator but may be of a more general type, the quasi-differential operator.[8] This new class of operators is proved to be closed with respect to taking closures and adjoints, and is therefore an operatorially correct generalization of the class of ordinary differential operators.

In §2 these results are extended to semi-infinite and infinite intervals by means of the important Lemma 2.2 on differential forms, which also has some interest of its own.

§3 deals with expressions of the form

$$p_{n_1, \cdots, n_t}(x_1, \cdots, x_t) \frac{\partial^{n_1 + \cdots + n_t}}{\partial x_1^{n_1} \cdots \partial x_t^{n_t}} + \sum_{\substack{r_k < n_k \\ k=1,\cdots,t}} p_{r_1, \cdots, r_t}(x_1, \cdots, x_t) \frac{\partial^{r_1 + \cdots + r_t}}{\partial x_1^{r_1} \cdots \partial x_t^{r_t}}.$$

---

[6] See for example K. Friedrichs: Mathematische Annalen, vol. 109, (1934), pp. 465–487 and 685–713. Also F. J. Murray: Transactions of the American Mathematical Society 37 (1935) pp. 301–338, who also gives other references.

[7] See E. L. Ince: *Ordinary Differential Equations*, Longmans, Green and Co. Ltd., London, 1927, pp. 123–124.

[8] See M. Bôcher: *Leçons sur les Méthodes de Sturm*, Gauthier-Villars et Cie, Paris, 1917, pp. 36–38.

Results are obtained which are analogous to those for the ordinary differential operator but we are led to an extension of the notion of linear boundary conditions, which, in integrated form, involves integration along the characteristic lines.

§4 indicates the essentially different behavior to be expected in the case of the elliptic operator.

The writer is greatly indebted to Professor J. von Neumann for proposing this problem and for inspiring guidance. In particular, Professor von Neumann supplied the proof of Lemma 2.1 on which is based the Lemma 2.2.

1. The Ordinary Differential Operator on a Finite Interval

In this section, H will denote the Hilbert space of the functions which are L.s.s. on a finite interval, which we take as $(0, 1)$ without loss of generality. The inner product (f, g) is given by the absolutely convergent integral $\int_0^1 f(x)\overline{g(x)}\,dx$, where $f(x)$ is any one of the class of Lebesgue equivalent functions determining the element f of H. When there is no possibility of confusion, the function $f(x)$ and the element f of H which it determines will be used interchangeably, but it is to be remembered that $f = g$ in H implies only $f(x) \sim g(x)$.[9]

Two linear sub-spaces D, D_0, dense in H, will be important as domains for operators which we shall consider. They are defined by

Definition 1.1. D is the linear space of the $f(x)$ for which
(i) $f^{(n-1)}(x)$ is abs. cont.
(ii) $f^{(n)}(x)$ is L.s.s.[10]

Definition 1.2. D_0 is the linear space of those $f(x)$ in D for which $f(x)$ $f'(x), \cdots, f^{(n-1)}(x)$, vanish for $x = 0$ and for $x = 1$.

Farther on we shall define and consider intermediate spaces D_ψ with $D \supset D_\psi \supset D_0$.

In the expression

$$E \equiv p_n(x)\frac{d^n}{dx^n} + p_{n-1}(x)\frac{d^{n-1}}{dx^{n-1}} + \cdots + p_0(x)$$

it will be assumed that
(i) $p_r^{(r-1)}(x)$ is abs. cont. for $1 \leqq r \leqq n$.
(ii) $p_r^{(r)}(x)$ is L.s.s. for $0 \leqq r \leqq n$.[10]
(iii) $|p_n(x)| \geqq \epsilon > 0$.

[9] $\sim$ here denotes equivalence in the sense of Lebesgue, that is, equality with the possible exception of a set of points of Lebesgue measure 0.

[10] Absolute continuity of $f^{(n-1)}(x)$ on the open interval $(0, 1)$ (that is, on every closed sub-interval) is here equivalent to absolute continuity on the closed interval $(0, 1)$, after proper definition of $f^{(n-1)}(0)$ and $f^{(n-1)}(1)$, due to the square summability of $f^{(n)}(x)$. We note also that the fundamental theorem of the (Lebesgue) integral calculus shows that the absolute continuity of $f^{(n-1)}(x)$ implies that of $f(x), f'(x), \cdots, f^{(n-2)}(x)$.

E will define, in the obvious way, linear operators T, T_0, with domains D, D_0, resp. We note that if the domain of T (resp. T_0) were restricted to the polynomials in D (resp. D_0), the closure, and hence the adjoint, would not be changed, (see footnote (5)). For to every f in D (resp. D_0) there is a sequence of polynomials $P_m(x)$ in D (resp. D_0) such that

$$\|f^{(n)}(x) - P_m^{(n)}(x)\| \to 0 \quad \text{as} \quad m \to \infty$$

and consequently the $P_m(x)$ could be chosen so that

$$P_m \to f \qquad \text{and} \qquad TP_m \to Tf \text{ as } m \to \infty.$$

Theorem 1.1. *The operators T and T_0 are closed; the adjoint operators are given by the Lagrange-Adjoint $\sum_{r=0}^{n} (-1)^r \dfrac{d^r(\overline{p_r(x)})}{dx^r}$ on the domains D_0, resp. D.*

Proof. Let P be the linear closed sub-space of H consisting of the polynomials of maximal degree $n - 1$. $f(x)$ will be in D_0 if and only if

$$f(x) = \int_0^x \frac{(x-z)^{n-1}}{(n-1)!} \varphi(z)\, dz$$

for some φ in $\ominus P$.[11] The statement $T_0^* g = g^*$, that is, g, g^*, are L.s.s., and $\int_0^1 (\sum_{r=0}^{n} p_r(x) f^{(r)}(x))\overline{g(x)}\, dx = \int_0^1 f(x)\overline{g^*(x)}\, dx$ for all f in D_0, can be written[12]

$$\int_0^1 p_n(z)\varphi(z)\overline{g(z)}\, dz + \sum_{r=0}^{n-1} \int_0^1 dz \int_0^z \frac{(z-x)^{n-r-1}}{(n-r-1)!} \varphi(x) p_r(z)\overline{g(z)}\, dx$$
$$= \int_0^1 dz \int_0^z \frac{(z-x)^{n-1}}{(n-1)!} \varphi(x)\overline{g^*(z)}\, dx.$$

Inverting the order of integration, we obtain

$$\int_0^1 p_n(z)\varphi(z)\overline{g(z)}\, dz + \sum_{r=0}^{n-1} \int_0^1 \varphi(x)\, dx \int_x^1 \frac{(z-x)^{n-r-1}}{(n-r-1)!} p_r(z)\overline{g(z)}\, dz$$
$$= \int_0^1 \varphi(x)\, dx \int_x^1 \frac{(z-x)^{n-1}}{(n-1)!} \overline{g^*(z)}\, dz.$$

---

[11] Every $f(x)$ in D clearly has the form

$$f(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + \int_0^x \frac{(x-z)^{n-1}}{(n-1)!} \varphi(z)\, dz, \qquad \varphi = f^{(n)},$$

(n as in Definition 1.1 and as in E!) for some constants $a_0, a_1, \cdots, a_{n-1}$. If $f(x)$ is even in D_0, then all $a_r = 0$, and $\int_0^1 (1-z)^{n-r-1} \varphi(z)dz = 0$ for $r < n$ (which is equivalent to $\int_0^1 z^r \varphi(z)dz = 0$ for $r < n$). This means φ is in $\ominus P$ (see footnote (4)).

[12] The changes in the order of the summations and integrations can easily be justified by use of Fubini's theorem, because of the restrictions on the $p_r(x)$, $f(x)$, $g(x)$, $g^*(x)$.

Written in the form

$$\int_0^1 \varphi(x)\left\{p_n(x)\overline{g(x)} + \sum_{r=0}^{n-1}\int_x^1 \frac{(z-x)^{n-r-1}}{(n-r-1)!}\,p_r(z)\overline{g(z)}\,dz - \int_x^1 \frac{(z-x)^{n-1}}{(n-1)!}\,\overline{g^*(z)}\,dz\right\}dx = 0$$

(for arbitrary φ in $\ominus P$!) this is the statement that

$$p_n(x)\overline{g(x)} + \sum_{r=0}^{n-1}\int_x^1 \frac{(z-x)^{n-r-1}}{(n-r-1)!}\,p_r(z)\overline{g(z)}\,dz - \int_x^1 \frac{(x-x)^{n-1}}{(n-1)!}\,\overline{g^*(z)}\,dz$$

is orthogonal to $\ominus P$, hence in P,[13] and thus

$$(*) \qquad \sim a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$$

for some constants $a_0, a_1, \cdots, a_{n-1}$.

The equivalence (*) can be made an equality by changing $g(x)$ on a null set since $p_n(x) \neq 0$. Then $p_n(x)\overline{g(x)}$, and with it $g(x)$, will be abs. cont. Differentiating,[14]

$$\frac{d}{dx}(p_n(x)\overline{g(x)}) - p_{n-1}(x)\overline{g(x)} - \sum_{r=0}^{n-2}\int_x^1 \frac{(z-x)^{n-r-2}}{(n-r-2)!}\,p_r(z)\overline{g(z)}\,dz + \int_x^1 \frac{(z-x)^{n-2}}{(n-2)!}\,\overline{g^*(z)}\,dz \sim a_1 + 2a_2 x + \cdots + (n-1)\,a_{n-1}x^{n-2}.$$

This shows that $\frac{d}{dx}(p_n(x)\overline{g(x)})$ is equivalent to an abs. cont. function and therefore abs. cont. without further change.[15] Repeated differentiation shows that, $T_0 g = g^*$ is equivalent to the conditions

(i) g is in D.

(ii) $g^*(x) \sim \sum_{r=0}^{n} (-1)^r \frac{d^r}{dx^r}(\overline{p_r(x)}g(x))$ and is L.s.s.

But if $g(x)$ is in D, $\sum_{r=0}^{n} (-1)^r \frac{d^r}{dx^r}(\overline{p_r(x)}g(x))$ is defined almost everywhere and is L.s.s. This proves the part of the theorem about the adjoint of T_0.

$T^* \subset T_0^*$ since $T \supset T_0$, so $T^*g = g^*$ implies that g is in D and $g^*(x) \sim \sum_{r=0}^{n}(-1)^r \frac{d^r(\overline{p_r(x)}g(x))}{dx^r}$. Integating by parts, the condition $\int_0^1 Tf\bar{g} = \int_0^1 f\overline{g^*}$ for all f in D, can be written as an addition condition

(iii) $[B(f(x), g(x))]_0^1 = 0$ for all f in D

where $B(f(x), g(x))$ is the bilinear concomitant, a form linear and homogeneous

---

[13] See footnote (4).

[14] If the differentiability conditions were not satisfied by the $p_r(x)$, repeated differentiation would lead to a quasi-differential operator (see footnote (8)).

[15] The theorem of Lebesgue implies that the derivative function of an abs. cont. function can itself be equivalent to an abs. cont. function only if it is abs. cont.

in $f(x), \cdots, f^{(n-1)}(x)$ and in $\overline{g(x)}, \cdots, \overline{g^{(n-1)}(x)}$, and with coefficients depending on the $p_r(x)$.[7] The determinant of the form $= \pm(p_n(x))^n \neq 0$. Since the $f(0), \cdots, f^{(n-1)}(0), f(1), \cdots, f^{(n-1)}(1)$ of the elements of D, are arbitrary complex numbers, g must be in D_0. This proves the part of the theorem concerning T^*.

The preceding paragraphs apply also to the expression

$$\sum_{r=0}^{n} p_r^*(x) \frac{d^r}{dx^r} \equiv \sum_{r=0}^{n} (-1)^r \frac{d^r}{dx^r} (\overline{p_r(x)})$$

since the

$$p_r^*(x) \equiv (-1)^n \binom{n}{r} \frac{d^{n-r}}{dx^{n-r}} \overline{p_n(x)} + (-1)^{n-1} \binom{n-1}{r-1} \frac{d^{n-r-1}}{dx^{n-r-1}} \overline{p_{n-1}(x)} + \cdots + (-1)^r \overline{p_r(x)}$$

will have $\dfrac{d^{r-1} p_r^*(x)}{dx^{r-1}}$ abs. cont., $\dfrac{d^r p_r^*(x)}{dx^r}$ L.s.s., and $|p_n^*(x)| \geqq \epsilon > 0$, due to the restrictions imposed on the $p_r(x)$. Hence T_0^{**} and T^{**} are single-valued operators with domains D_0 and D resp. Since $T_0 \subset \tilde{T}_0 = T_0^{**}$ and $T \subset \tilde{T} = T^{**}$ it follows that $T_0 = \tilde{T}_0$ and $T = \tilde{T}$. This completes the proof of the theorem. (We note that this gives a proof of $p_r^{**}(x) = p_r(x)$.[7] A direct proof of $T = \tilde{T}$, $T_0 = \tilde{T}_0$ can be obtained from Lemma 2.2.)

By linear boundary conditions we can define a domain D_ψ, $D \supset D_\psi \supset D_0$, and an operator T_ψ as follows:

DEFINITION 1.3. D_ψ is the linear space of the $f(x)$ in D which satisfy

$$\psi_k\{f(0), \cdots, f^{(n-1)}(0), f(1), \cdots, f^{(n-1)}(1)\} = 0$$
$$k = 1, \cdots, m \qquad m \leqq 2n$$

where the ψ_k are linear, homogeneous and independent.

The matrix of the ψ_k will be of rank m and we may assume without loss of generality that the minor of order m on the left is non-singular. We can then write the conditions in the matrix form

$$\left\| \begin{matrix} f(0) \\ f'(0) \\ \cdot \\ \cdot \\ \cdot \\ f^{(n-1)}(1) \end{matrix} \right\| = \left\| \begin{matrix} 1 \\ \\ \\ M \end{matrix} \right\| \left\| \begin{matrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_{2n-m} \end{matrix} \right\|$$

where M is a matrix of m rows and $2n - m$ columns, 1 is the unit matrix of $2n - m$ rows and columns, and $u_1, \cdots, u_{2n-m}$ may take arbitrary complex values. The expression E will determine an operator T_ψ on the domain D_ψ.

THEOREM 1.2. *T_ψ is a closed linear operator satisfying $T \supset T_\psi \supset T_0$ and every closed linear operator intermediate to T and T_0 can be put in this form; the corresponding adjoint is given by the Lagrange-Adjoint on a linear sub-space of D defined by an adjoint set of boundary conditions.*

PROOF. Since $T_0 \subset T_\psi$ we obtain, as in the proof of Theorem 1.1, that $T_\psi^* g = g^*$ is equivalent to the three conditions

(i) g lies in D,

(ii) $g^* = \sum_{r=0}^{n} (-1)^r \dfrac{d^r}{dx^r} (\bar{p}_r g)$,

(iii) $[B(f(x), g(x))]_0^1 = 0$ for all f in D_ψ.

The last condition can be written

$$\begin{Vmatrix} \overline{g(0)} & \overline{g'(0)} & \cdots & \overline{g^{(n-1)}(1)} \end{Vmatrix} \begin{Vmatrix} B_0 & 0 \\ 0 & -B_1 \end{Vmatrix} \begin{Vmatrix} f(0) \\ f'(0) \\ \cdot \\ \cdot \\ \cdot \\ f^{(n-1)}(1) \end{Vmatrix} = 0$$

where B_0, B_1 are the matrices of B for $x = 0$, $x = 1$, respectively; that is

$$\begin{Vmatrix} \overline{g(0)} & \overline{g'(0)} & \cdots & \overline{g^{(n-1)}(1)} \end{Vmatrix} \begin{Vmatrix} B_0 & 0 \\ 0 & -B_1 \end{Vmatrix} \begin{Vmatrix} 1 \\ M \end{Vmatrix} \begin{Vmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_{2n-m} \end{Vmatrix} = 0$$

for all $u_1, \cdots u_{2n-m}$. Transposing and taking complex-conjugate values, we obtain

$$\begin{Vmatrix} 1 & M^* \end{Vmatrix} \begin{Vmatrix} B_0^* & 0 \\ 0 & -B_1^* \end{Vmatrix} \begin{Vmatrix} g(0) \\ g'(0) \\ \cdot \\ \cdot \\ \cdot \\ g^{(n-1)}(1) \end{Vmatrix} = 0.$$

Since B_0^*, B_1^* are non-singular, this gives $2n - m$ linearly independent conditions

$$\psi_k^* \{g(0), g'(0), \cdots g^{(n-1)}(1)\} = 0 \qquad k = 1, \cdots, 2n - m$$

defining a linear sub-space D_ψ^* of D, which we have now shown to be the domain of T_ψ^*.

In the same way the closure $\tilde{T}_\psi = T_\psi^{**}$ will have a domain D_ψ^{**} which contains D_ψ and is determined by the same number $2n - (2n - m) = m$ of linearly independent boundary conditions. Thus $D_\psi = D_\psi^{**}$ and $T_\psi = \tilde{T}_\psi$ is a closed operator.

Now let T_c be any closed linear operator intermediate to T_0 and T, and let the ordered sets of values $(f(0), \cdots, f^{(n-1)}(0), f(1), \cdots, f^{(n-1)}(1))$ assumed by the $f(x)$ in the domain of T_c, form a linear sub-set of $2n$-dimensional unitary space[16] defined by the conditions

$$\psi_k\{f(0), \cdots, f^{(n-1)}(0), f(1), \cdots, f^{(n-1)}(1)\} = 0 \qquad k = 1, 2, \cdots, m.$$

Let T_ψ be the operator determined by these conditions as on page 886. Then $T_c \subset T_\psi$ and $T_c^* \supset T_\psi^*$. But for fixed g the vanishing of $[B(f(x), g(x))]_0^1$ (for all f in the domain of T_c!) depends only on the possible values of $f(0), \cdots, f^{(n-1)}(1)$. Consequently $T_c^* = T_\psi^*$ and $T_c = T_c^{**} = T_\psi^{**} = T_\psi$. Thus all parts of the theorem are proved.

Examination of the above determination of T^*, T_0^*, shows that if the $p_r(x)$ are not restricted, T^*, T_0^*, will not be differential operators but will be quasi-differential operators (see footnote 14), (these will be described below). Instead of attempting to weaken the restrictions on the coefficients $p_r(x)$, we consider at once the more general quasi-differential operators.

Let $q_0(x), \cdots, q_n(x)$, $p_0(x), \cdots, p_n(x)$ be measurable and bounded and $|q_0(x)|$, $|p_n(x)| \geqq \epsilon$ for some $\epsilon > 0$. Define the linear sub-spaces D', D_0' in H by the

Definition 1.1′. D' is the linear space of those L.s.s. functions $f(x)$ for which the equivalences

$$\begin{aligned} f_{(0)}(x) &\sim q_0(x)f(x) \\ f_{(1)}(x) &\sim f_{(0)}'(x) + q_1(x)f(x) \\ &\cdot \quad \cdot \quad \cdot \\ &\cdot \quad \cdot \quad \cdot \\ f_{(n)}(x) &\sim f_{(n-1)}'(x) + q_n(x)f(x) \end{aligned}$$

define functions $f_{(r)}(x)$ which are abs. cont. for $0 \leqq r < n$ and L.s.s. for $r = n$.

Remark. Let

$$K_0(x, z) = 1 - \int_z^x \frac{q_1(z_1)}{q_0(z_1)} dz_1 + \int_z^x \frac{q_1(z_1)}{q_0(z_1)} dz_1 \int_{z_1}^x \frac{q_1(z_2)}{q_0(z_2)} dz_2 - \cdots$$

and

$$\begin{aligned} K_r(x, z) = 1 &- \int_z^x dz_1 \int_{z_1}^x K_{r-1}(z_2, z_1) dz_2 \int_{z_2}^x K_{r-2} \cdots \int_{z_r}^x K_0(z_{r+1}, z_r) \frac{q_{r+1}(z_{r+1})}{q_0(z_{r+1})} dz_{r+1} \\ &+ \int_z^x dz_1 \int_{z_1}^x K_{r-1} \cdots \int_{z_r}^x K_0(z_{r+1}, z_r) \frac{q_{r+1}(z_{r+1})}{q_0(z_{r+1})} dz_{r+1} \int_{z_{r+1}}^x dz_{r+2} \int_{z_{r+2}}^x K_{r-1} \cdots \\ &\qquad \int_{z_{2r+1}}^x K_0(z_{2r+2}, z_{2r+1}) \frac{q_{r+1}(z_{2r+2})}{q_0(z_{2r+2})} dz_{2r+2} - \cdots \end{aligned}$$

$$0 < r < n.$$

[16] That is, the space of all $(x_1, \cdots, x_{2n})$, the x_i arbitrary complex numbers.

Then the kernels $K_r(x, z)$, $r = 0, 1, \cdots, n - 1$, have the following property: If $f(x)$ is in D', with $f_{(0)}(0) = f_{(1)}(0) = \cdots = f_{(n-1)}(0) = 0$, then

$$(\dagger) \qquad f(x) = \frac{f_{(0)}(x)}{q_0(x)}, \qquad f_{(r)}(x) = \int_0^x K_r(x, z) f_{(r+1)}(z)\, dz \qquad \text{for } 0 \leqq r < n.$$

(The infinite series for the kernels can be obtained in succession for $r = 0, 1, \cdots, n - 1$, from the relations $f_{(r)}(x) = \int_0^x f_{(r-1)}(z)\, dz - \int_0^x q_{r+1}(z) f(z)\, dz$ by the well-known methods of successive substitutions. The uniform convergence of these infinite series, the boundedness of the kernels, and the relations (†) can be obtained easily from the boundedness of the $q_r(x)/q_0(x)$. We shall omit the details.) Conversely, if $\varphi(x)$ is any L.s.s. function, then the equations (†) taken in the order $r = n - 1, \cdots, 1, 0$, and with $f_{(n)}(x)$ replaced by $\varphi(x)$, will define an $f(x)$ in D' such that $f_{(n)}(x) = \varphi(x)$ and $f_{(r)}(0) = 0$ for $r = 0, 1, \cdots, n - 1$. By the use of successive substitutions we can also define for every

$$s = 0, 1, \cdots, n - 1,$$

a (unique) function $M^s(x)$ in D' such that $M^s_{(n)}(x) = 0$ and $M^s_{(r)}(0) = 1$ for $r = s$ and $= 0$ for $r \neq s$. Then every $f(x)$ in D' has a unique representation in the form

$$\sum_{s=0}^{n-1} f_{(s)}(0) M^s(x) + \frac{1}{q_0(x)} \int_0^x K_0(x, z_1)\, dz_1 \int_0^{z_1} \cdots \int_0^{z_{n-1}} K_{n-1}(z_{n-1}, z_n) f_{(n)}(z_n)\, dz_n$$

where the $f_{(r)}(0)$ may be arbitrary numbers and the $f_{(n)}(x)$ may be any L.s.s. function.[17]

Definition 1.2′. D_0' is the linear space of those $f(x)$ in D' for which $f_{(0)}(x), \cdots, f_{(n-1)}(x)$ vanish for $x = 0$ and for $x = 1$.

Define the quasi-differential operators T', T_0' with domains D', D_0' resp., by $T' \supset T_0'$ and

$$T'f(x) = \sum_{r=0}^{n} p_r(x) f_{(r)}(x).$$

Let the domains D'', D_0'' and operators T'', T_0'', be defined in the same way when $q_r(x)$, $p_r(x)$ are replaced by $(-1)^{n-r}\,\overline{p_{n-r}(x)}$, $(-1)^{n-r}\,\overline{q_{n-r}(x)}$, resp.

Theorem 1.1′. *T', T_0' are single-valued closed operators with domains dense in H. The adjoint operators are T_0'', T'', resp.*

---

[17] The ordinary differentiation previously considered, is the special case of quasi-differentiation when $q_0(x) \equiv 1$, and $q_r(x) \equiv 0$ for $r \neq 0$; the $M^s(x)$ become polynomials of degree $n - 1$ in x and $K_r(x, z) \equiv 1$ for all r, x, z.

PROOF.[18] We will show that if $(T')^*g$ is defined at all, then g must be in D_0''. For if $(T')^*g$ has a value g^* then

$$\int_0^1 \left(\sum_{r=0}^{n} p_r(x) f_{(r)}(x) \right) \overline{g(x)}\, dx = \int_0^1 f(x) \overline{g^*(x)}\, dx$$

for every f in D', which implies that for every L.s.s. φ (use (†)),

$$\int^1 \varphi(x) \Big[p_n(x) \overline{g(x)} + \int_x^1 K_{n-1}(z, x) p_{n-1}(z) \overline{g(z)}\, dz + \cdots$$

$$+ \int_x^1 K_{n-1} \int_{z_1}^1 \cdots \int_{z_{n-1}}^1 K_0(z_{n-1}, z_{n-2}) \left(\overline{g(z_{n-1})} p_0(z_{n-1}) - \frac{\overline{g^*(z_{n-1})}}{q_0(z_{n-1})} \right) dz_{n-1} \Big] dx = 0.$$

Thus the expression $[\cdots]$ in the preceding line is equivalent to zero. Repeated differentiation, which can be justified by using the explicit formulae for $K_r(x, z)$ and the restrictions on the $q_r(x)$, $p_r(x)$, shows that g must be in D''. If we differentiate $[\cdots] = 0$ r times and then set $x = 1$ we obtain $g_{(r)}(1) = 0$ for $0 \leqq r < n$. Interchanging the rôles of the end-points of the interval (0, 1) we obtain $g_{(r)}(0) = 0$ for $0 \leqq r < n$. Thus g is in D_0'', proving that the domain of $(T')^*$ lies in D_0''.

It will be shown in the Remark to Lemma 2.2, page 899, that T', T_0' are closed. Since T' is single-valued and is the adjoint of $(T')^*$, it follows that $(T')^*$ has a domain dense in H. Thus D_0'', and a fortiori D'', are dense in H. It can be shown in the same way that D_0', D' are dense in H.

Thus T', T_0' are single-valued closed linear operators with domains dense in H and their adjoints will be single-valued with domains contained in D_0'', D'', resp. Theorem 1.1′ now follows from the identity

$$(\#) \qquad \int_0^1 (T'f(x) \cdot \overline{g(x)} - f(x) \cdot \overline{T''g(x)})\, dx = \sum_{r=0}^{n-1} (-1)^r [f_{(n-r-1)}(x)\, \overline{g_{(r)}(x)}]_0^1$$

for all f in D', and g in D''. The identity (#) can be verified by integration by parts.

With the help of (#), linear boundary conditions on the $f_{(0)}(x), \cdots, f_{(n-1)}(x)$ could be discussed,[19] to give a generalization of Theorem 1.2. We may also

---

[18] A proof on the same lines as the proof of Theorem 1.1, would require a number of identities in the $K_r(x, z)$ and their partial differential coefficients. These identities, although not particularly difficult to establish, involve many long formulae, and are therefore avoided here.

[19] It would be necessary to prove that all (arbitrary) sets of $2n$ numbers are assumed as values by the $(f_{(0)}(0), \cdots, f_{(n-1)}(0), f_{(0)}(1), \cdots, f_{(n-1)}(1))$ when f runs over D'. Let $\bar{D}'$ be the set of f's in D' for which $f_{(0)}(0) = f_{(1)}(0) = \cdots = f_{(n-1)}(0) = 0$; then it is sufficient to show that $(f_{(0)}(1), \cdots, f_{(n-1)}(1))$ takes on all arbitrary sets of n numbers as values when f runs over $\bar{D}'$. This may be shown as follows: Let $\bar{T}'$ be defined with domain $\bar{D}'$ by $\bar{T}' \subset T'$. Then the proof for Theorem 1.1′ shows that if g is in the domain of $(\bar{T}')^*$, then necessarily

remark that the results of this section extend with little difficulty to operational expressions of the form $\sum_{r=0}^{n} p_r(x_1, \cdots, x_l) \frac{\partial^r}{\partial x_1^r}$.

2. The Infinite Interval

The results of section 1 can be extended to infinite and semi-infinite intervals by means of the following Lemma 2.2 on differential forms. The lemma will first be formulated for the special case of ordinary differential forms and the reader may find it advantageous to consider the proof given after first simplifying it for this special case: the proof will be given here at once for the more general quasi-differential forms.

Lemma 2.1.[20] *Let $f(x)$ be defined on a finite or infinite interval (a, b). Let $p_r(x)$, $r = 0, 1, \cdots, n$, be measurable on (a, b) with $|p_r(x)| \leqq C < \infty$ for $0 \leqq r < n$, and $|p_n(x)| \geqq \epsilon > 0$. If*

(i) *$f^{(n-1)}(x)$ is abs. cont. on every (c, d), $a < c < d < b$,*

(ii) *$f(x)$ and $\sum_{r=0}^{n} p_r(x) f^{(r)}(x)$ are L.s.s. on (a, b), then $f'(x), \cdots, f^{(n)}(x)$ are all L.s.s. on (a, b).*

Lemma 2.2. *Let $f(x)$ be defined on a finite or infinite interval (a, b). Let $q_0(x), \cdots, q_n(x)$, $p_0(x), \cdots, p_n(x)$ be measurable and $|q_r(x)|$, $|p_{n-r}(x)|$, $\leqq C < \infty$ for $r = 1, \cdots, n$, and $|q_0(x)|$, $|p_n(x)|$, $\geqq \epsilon > 0$, on (a, b). Assume that the equivalences*

$$f_{(0)}(x) \sim q_0(x) f(x)$$
$$f_{(1)}(x) \sim f'_{(0)}(x) + q_1(x) f(x)$$
$$\cdot \quad \cdot \quad \cdot$$
$$\cdot \quad \cdot \quad \cdot$$
$$f_{(n)}(x) \sim f^1_{(n-1)}(x) + q_n(x) f(x)$$

$g_{(r)}(1) = 0$ for $r = 0, 1, \cdots, n-1$. On the other hand, the identity (✱) shows that g will be in the domain of $(\bar{T}')^*$ if only

(✱✱) g is in D'', and $\sum_{r=0}^{n-1} (-1)^r \overline{g_r(1)} f_{(n-r-1)}(1) = 0$ for all $(f_{(r)}(1); r = 0, 1, \cdots, n-1)$ when f runs over $\bar{D}'$.

Now the corollary to Definition 1.1, with $q_r(x)$ replaced by $(-1)^{n-r}\overline{p_{n-r}(x)}$ and the rôles of the end-points of the interval $(0, 1)$ interchanged, shows that $(g_{(r)}(1); r = 0, 1, \cdots, n-1)$ assumes all arbitrary sets of n numbers as values when g runs over D''. Since (✱✱) implies $g_{(r)}(1) = 0$ for $r = 0, 1, \cdots, n-1$, this means that $(f_{(r)}(1); r = 0, 1, \cdots, n-1)$ assumes all arbitrary sets of n numbers as values when f runs over $\bar{D}'$.

[20] The proof of Lemma 2.1 was found by J. von Neumann; the details of his proof are essential in the proof of the more general Lemma 2.2. The lemma can be proved also in several variables (see Lemma 3.1). A theorem of the type of Lemma 2.1 (with 'L.s.s.' replaced throughout by 'bounded') was first given, but with differentiability restrictions on the $p_r(x)$, by E. Esclangon. Esclangon's theorem, without the restrictions, was shown by E. Landau (Mathematische Annalen 102 (1929) pp. 177–78) to be an easy consequence of a simple lemma of Hardy-Littlewood. We may remark that Esclangon's theorem can also be proved for the more general quasi-differential expression.

define $f_{(r)}(x)$ *abs. cont. on every* (c, d), $a < c < d < b$, *for* $0 \leqq r < n$, *and (therefore necessarily)* $f_{(n)}(x)$ *almost everywhere on* (a, b). *If* $f(x)$ *and* $\sum_{r=0}^{n} p_r(x)f_{(r)}(x)$ *are L.s.s. on* (a, b), *then* $f_{(r)}(x)$ *will be L.s.s. on* (a, b), *for* $0 \leqq r \leqq n$.

PROOF. .1 We may (and shall) assume that the interval is of the type (0, b), that $f_{(r)}(0) = 0$ for $0 \leqq r < n$, and that $q_0(x)$, $p_n(x)$ are identically equal to 1. For we may clearly assume $a < 0 < b$, and it is sufficient to prove the lemma for each of $(a, 0)$, $(0, b)$ separately. We need consider only $(0, b)$ since the $(a, 0)$ can be transformed into $(0, -a)$ by replacing $f(x)$ by $f(-x)$, and $q_r(x)$, $p_r(x)$ by $(-1)^r q_r(-x)$, $(-1)^r p_r(-x)$ resp. Now define a polynomial $P(x)$ such that

$$P^{(r)}(0) = f_{(r)}(0) \qquad 0 \leqq r < n,$$
$$P^{(r)}(-1) = 0 \qquad 0 \leqq r < n;$$

this can clearly be done. The function $h(x)$,

$$h(x) \begin{cases} = f(x) & 0 \leqq x < b, \\ = P(x) & -1 \leqq x < 0, \end{cases}$$

will satisfy the hypotheses of the lemma on the interval $(-1, b)$, if we set $q_r(x)$, $p_{n-r}(x)$, $= 0$ for $1 \leqq r \leqq n$, and $q_0(x)$, $p_n(x)$, $= \epsilon$ for $-1 \leqq x < 0$, and it is sufficient to prove the lemma for such an $h(x)$ on $(-1, b)$. Replacing x by $x + 1$, we see that we may restrict ourselves to the case of an interval $(0, b)$ and an $f(x)$ for which $f_{(0)}(0) = \cdots = f_{(n-1)}(0) = 0$. We may also assume $q_0(x)$, $p_n(x)$, $\equiv 1$ by replacing $f(x)$, $q_r(x)$, $p_r(x)$, C by $q_0(x)f(x)$, $\dfrac{q_r(x)}{q_0(x)}$, $\dfrac{p_r(x)}{p_n(x)}$, $\dfrac{C}{\epsilon}$, resp.

.2 Let

$$Lf(x) = f_{(n)}(x) + \sum_{r=0}^{n-1} p_r(x)f_{(r)}(x) \qquad W(x) = \sum_{r=0}^{n-1} |f_{(r)}(x)|^2.$$

Using the inequality of Schwarz,

$$|f_{(n)}(x)|^2 \leqq \{|Lf(x)| + \sum_{r=0}^{n-1} C\,|f_{(r)}(x)|\}^2 \leqq (n+1)\,\{|Lf(x)|^2 + C^2 W(x)\},$$

$$\text{(I)} \qquad \int_0^b |f_{(n)}(x)|^2\,dx \leqq (n+1)\left\{\int_0^b |Lf(x)|^2\,dx + \int_0^b W(x)\,dx\right\}.$$

Thus $\int_0^b |f_{(n)}(x)|^2\,dx$ will be finite if we can show that $\int_0^b W(x)\,dx$ is finite; the same is true for every $\int_0^b |f_{(r)}(x)|^2\,dx$, $0 \leqq r < n$, from the definition of $W(x)$. Hence we need only prove $\int_0^b W(x)\,dx$ finite to prove the lemma.

.3 We shall now derive an inequality for $W(x)$ which will show that $\int_0^b W(x)\,dx$ is finite if b is finite. In the case of b infinite, and assuming that

$\int_0^b W(x)\,dx$ is infinite, we can derive a second inequality which contradicts the first one, showing that $\int_0^b W(x)\,dx$ is always finite, and so proving the lemma. We proceed to derive these inequalities.

.4 For $0 \leqq x < b$ we have

$$\begin{aligned}
|W'(x)| &\leqq \sum_{r=0}^{n-1} 2\,|f_{(r)}(x)|\,|f'_{(r)}(x)| \\
&= \sum_{r=0}^{n-1} 2\,|f_{(r)}(x)|\,|f_{(r+1)}(x) - q_{r+1}(x)f(x)| \\
&\leqq \sum_{r=0}^{n-1} \{|f_{(r)}(x)|^2 + |f_{(r+1)}(x) - q_{r+1}(x)f(x)|^2\} \\
&\leqq \sum_{r=0}^{n-1} \{|f_{(r)}(x)|^2 + 2\,|f_{(r+1)}(x)|^2 + 2C^2|f(x)|^2\} \qquad \text{(use (I) in .2)} \\
&\leqq 3W(x) + 2(n+1)\{|Lf(x)|^2 + C^2W(x)\} + 2nC^2|f(x)|^2 \\
&\leqq \{3 + 4(n+1)C^2\}\,W(x) + 4n\,|Lf(x)|^2.
\end{aligned}$$

Integrating this differential inequality, and remembering that $W(0) = 0$ because of the restrictions on $f_{(r)}(0)$, as described in .1, we obtain

$$W(x) = \int_0^x W'(x_1)\,dx_1 \leqq \int_0^x |W'(x_1)|\,dx_1 \leqq \{3 + 4(n+1)C^2\}\int_0^x W(x_1)\,dx_1 + 4n\int_0^x |Lf(x_1)|^2\,dx_1.$$

If we set $A = \{3 + 4(n+1)C^2\}$ and $B = 4n \int_0^b |Lf(x)|^2\,dx$, then

$$W(x) \leqq B + A\int_0^x W(x_1)\,dx_1,$$

$$\frac{d}{dx}\left\{e^{-Ax}\int_0^x W(x_1)\,dx_1\right\} \leqq Be^{-Ax},$$

$$\int_0^x W(x_1)\,dx_1 = e^{Ax}\int_0^x \left[\frac{d}{dx_1}\left\{e^{-Ax_1}\int_0^{x_1} W(x_2)\,dx_2\right\}\right]dx_1 \leqq e^{Ax}\int_0^x Be^{-Ax_1}\,dx_1 = e^{Ax}\frac{B}{A}(1 - e^{-Ax}) = \frac{B}{A}(e^{Ax} - 1).$$

Hence

$$W(x) \leqq B + A\int_0^x W(x_1)\,dx_1 \leqq B + A\frac{B}{A}(e^{Ax} - 1) = Be^{Ax},$$

giving the first of the two inequalities mentioned in .3,

(1) $$W(x) \leqq Be^{Ax} \qquad \text{for } 0 \leqq x \leqq b.$$

(1) makes it clear that $\int_0^b W(x_1)\,dx_1$ is finite if b is finite.

.5 For $0 \leqq x < b$, we have

$$\begin{aligned}\{(|f_{(r-1)}|^2)' + 2\Re(\overline{f_{(r-1)}}q_r f)\}' &= 2\Re(\overline{f_{(r-1)}}f_{(r)})' \\ &= 2\Re\{|f_{(r)}|^2 - \bar{q}_r\bar{f}f_{(r)} + \overline{f_{(r-1)}}f_{(r+1)} - q_{r+1}\overline{f_{(r-1)}}f\} \\ &= 2\Re\{|f_{(r)}|^2 - \bar{f}(\bar{q}_r f_{(r)} + \overline{q_{r+1}}f_{(r-1)}) + \overline{f_{(r-1)}}f_{(r+1)}\} \\ &\geqq 2|f_{(r)}|^2 - |f_{(r-1)}|^2 - |f_{(r+1)}|^2 - C\{2|f|^2 + |f_{(r)}|^2 + |f_{(r-1)}|^2\},\end{aligned}$$

$$\begin{aligned}(2 - C)|f_{(r)}|^2 \leqq (1 + C)|f_{(r-1)}|^2 + |f_{(r+1)}|^2 + 2C|f|^2 \\ + \{(|f_{(r-1)}|^2)' + 2\Re(\overline{f_{(r-1)}}q_r f)\}',\end{aligned}$$

and hence

(Δ) $$|f_{(r)}|^2 \leqq C_1\{|f_{(r-1)}|^2 + |f_{(r+1)}|^2\} + C_2|f|^2 + C_3\{(|f_{(r-1)}|^2)' + 2\Re(\overline{f_{(r-1)}}q_r f)\}'$$

where C_1, C_2, C_3 may be taken as constants (that is, independent of r and x) which depend on C and converge to the limits $\frac{1}{2}$, 1, $\frac{1}{2}$, resp. when C tends to zero. Since the results of the present paragraph will be applied only for the case of small C (namely in .6, where this case is obtained by a transformation $y = \theta x$), we shall restrict ourselves for the remainder of this paragraph to the case of small C. Now (Δ), for $r = 1, \cdots, n - 1$, is a convexity-like inequality in $|f_{(r-1)}|^2$, $|f_{(r)}|^2$, $|f_{(r+1)}|^2$; if C is sufficiently small, and hence C_1 sufficiently close to $\frac{1}{2}$, then the set of inequalities (Δ) will imply, for $r = 1, \cdots, n - 1$,[21]

[21] That (Δ) implies ($\Delta\Delta$), can be shown by the following discussion of inequalities: Let $v, \lambda_0, \lambda_1, \cdots, \lambda_n$ be non-negative numbers (with $|v - \frac{1}{2}| <$ a certain ϵ which will be made precise shortly), and let $\mu_1, \cdots, \mu_{n-1}$ be any numbers such that the λ_r, μ_r satisfy the convexity-like inequalities

(Δ') $$\lambda_R \leqq v(\lambda_{R-1} + \lambda_{R+1}) + \mu_R \qquad R = 1, \cdots, n - 1$$

Define real numbers $c_0, c_1, \cdots, c_{n-1}$ by

(c) $$c_0 = 0, \quad c_1 = 1, \quad \text{and} \quad c_r = \frac{c_{r-1}}{v} - c_{r-2} \qquad \text{for } r = 2, \cdots, n - 1$$

and set $w_r = c_r c_{n-r} - v(c_r c_{n-r-1} + c_{r-1}c_{n-r})$ for $r = 1, \cdots, n - 1$. If $v = \frac{1}{2}$ then $c_r = r$, ($c_r = r$ satisfy (c), and (c) clearly has a unique solution), and $w_r = \frac{1}{2}n$. Since the c_r are continuous in v, there is an $\epsilon > 0$ such that $|v - \frac{1}{2}| < \epsilon$ implies $1/\epsilon > c_r > 0$ and $|w_r| > \epsilon > 0$ for $r = 1, \cdots, n - 1$. Now assume that v is such that we do have $|v - \frac{1}{2}|$ less than this ϵ.

Multiply both sides of (Δ') by $c_{\min(r,R)}c_{\min(n-r,n-R)}$, r being fixed and $= 1, \cdots, n - 1$, and sum over $R = 1, \cdots, n - 1$. Because of (c) the result is

$$w_r\lambda_r \leqq vc_{n-r}\lambda_0 + vc_r\lambda_n + \sum_{R=1}^{n-1} u_R\mu_R$$

$$(\Delta\Delta)\quad |f_{(r)}|^2 \leqq h_1(C,r)|f|^2 + h_2(C,r)|f_{(n)}|^2 + h_3(C,r)|f|^2 + \sum_{R=1}^{n-1} a_R(C,r)\{(|f_{(R-1)}|^2)' + 2\Re(\overline{f_{(R-1)}}\,q_r f)\}'$$

where the $h_1(C, r)$, $h_2(C, r)$, $|a_R(C, r)|$ are bounded for all sufficiently small C. Adding the inequalities $(\Delta\Delta)$ for $r = 1, \cdots, n - 1$, together with $|f|^2 = |f|^2$, gives (use (I) in .2),

$$W(x) \leqq h(C)|f(x)|^2 + k(C)\{|Lf(x)|^2 + C^2 W(x)\} + \sum_{R=1}^{n-1} a_R(C)\{(|f_{(R-1)}|^2)' + 2\Re(\overline{f_{(R-1)}}\,q_R f)\}'$$

where the $h(C)$, $k(C)$, $|a_R(C)|$ are all bounded, $\leqq K$ (say), for a finite K independent of the (sufficiently small) C. Integrating both sides of the last inequality gives (remember that $f_{(0)}(0) = 0$ and $(|f_{(r)}(x)|^2)'_{x=0} = 0$ for $r = 0, 1, \cdots, n - 1$),[22]

$$\int_0^x (1 - KC^2)W(x_1)\,dx_1 \leqq K\left\{\int_0^b |f(x_1)|^2 dx_1 + \int_0^b |Lf(x_1)|^2 dx_1 + 2\sum_{r=0}^{n-2} |f_{(r)}(x)\,q_{r+1}(x)f(x)|\right\} + \sum_{r=0}^{n-2} a_{r+1}(C)(|f_{(r)}(x)|^2)'$$

$$\leqq K\left\{\int_0^b |f|^2 + \int_0^b |Lf|^2\right\} + KC\sum_{r=0}^{n-1}\{|f(x)|^2 + |f_{(r)}(x)|^2\} + \sum_{r=0}^{n-2} a_{r+1}(C)(|f_{(r)}(x)|^2)' \leqq \text{(use (I) in .2)}$$

$$\leqq K\left\{\int_0^b |f|^2 + \int_0^b |Lf|^2 + nC|f(x)|^2\right\} + KCW(x) + \sum_{r=0}^{n-2} a_{r+1}(C)(|f_{(r)}(x)|^2)'.$$

Another integration gives (remember $|f_{(r)}(0)|^2 = 0$ for $r = 0, 1, \cdots, n - 1$)

$$(1 - KC^2)\int_0^x dx_1 \int_0^{x_1} W(x_2)\,dx_2 \leqq K\left\{\int_0^\infty |f|^2 + \int_0^\infty |Lf|^2\right\}\cdot x + nKC\int_0^b |f|^2 + KC\int_0^x W(x_1)\,dx_1 + KW(x).$$

---

where the $|u_R|$ are all less than some finite M independent of v. Dividing both sides of this inequality by w_r, $(w_r > \epsilon > 0)$, we obtain

$$(\Delta'\Delta')\qquad \lambda_r \leqq h_1(v, r)\lambda_0 + h_2(v, r)\lambda_n + \sum_{R=1}^{n-1} a_R(v, r)\mu_R$$

where the h_1, h_2, and $|a_R|$ are $< \overline{M}$, for some finite $\overline{M}$ independent of v. Thus (Δ') implies $(\Delta'\Delta')$, and in the same way, (Δ) implies $(\Delta\Delta)$.

[22] $(|f_{(r)}(x)|^2)'_{x=0} = (f_{(r)}(x)\overline{f_{(r)}(x)})'_{x=0} = (f'_{(r)}(x)\overline{f_{(r)}(x)} + f_{(r)}(x)\overline{f'_{(r)}(x)})_{x=0} = 0$

since $f_{(n)}(0) = 0$ for $r = 0, 1, \cdots, n - 1$, as secured in .1.

Since $1 - KC^2 > \frac{1}{2}$ for all sufficiently small C, we have

$$\text{(II)} \quad \int_0^x dx_1 \int_0^{x_1} W(x_2)\,dx_2 \leqq \bar{K}\left\{\int_0^b |f|^2 + \int_0^b |Lf|^2\right\}\cdot x + \bar{K}C\int_0^b |f|^2 + \bar{K}C\int_0^x W(x_1)\,dx_1 + \bar{K}W(x)$$

where $\bar{K} = 2nK$ is independent of the (sufficiently small) C.

.6 Since $W(x) \geqq 0$ for all x, $\int_0^b W(x)\,dx$ is finite or $= +\infty$. Assume, if possible, that $\int_0^b W(x)\,dx = +\infty$, and, consequently, that $b = +\infty$ $\left(.4\text{ shows that } \int_0^b W(x)\,dx \text{ is finite if } b \neq +\infty\right)$. Let $0 < \theta < 1$ and consider the functions $f^\theta(x) = f(\theta x)$, $q_r^\theta(x) = \theta^r q_r(\theta x)$, $p_r^\theta(x) = \theta^{n-r} p_r(\theta x)$.

$$q_0^\theta(x), \cdots, q_n^\theta(x), p_0^\theta(x), \cdots, p_n^\theta(x)$$

are measurable, and $|q_r^\theta(x)|, |p_{n-r}^\theta(x)|, \leqq C^\theta < \infty$ $(C^\theta = \theta C)$, for $r = 1, \cdots, n$, and $q_0^\theta(x)$, $p_n^\theta(x)$, are identically $= 1$. The equivalences

$$f_{(0)\theta}^\theta(x) \sim q_0^\theta(x) f^\theta(x)$$
$$f_{(1)\theta}^\theta(x) \sim (f_{(0)\theta}^\theta(x))' + q_1^\theta(x) f^\theta(x)$$
$$\cdot \quad \cdot \quad \cdot$$
$$\cdot \quad \cdot \quad \cdot$$
$$f_{(n)\theta}^\theta(x) \sim (f_{(n-1)\theta}^\theta(x))' + q_n^\theta(x) f^\theta(x)$$

define the functions $f_{(r)\theta}^\theta(x)$ $(= \theta^r f_{(r)}(\theta x))$ abs. cont. on every $(0, d)$, $d < \infty$, for $0 \leqq r < n$, and $f^\theta(x)$, $\sum_{r=0}^n p_r^\theta(x) f_{(r)\theta}^\theta(x)$ $(= \theta^n \sum_{r=0}^n p_r(\theta x) f_{(r)}(\theta x))$, are L.s.s. on $(0, \infty)$. Hence if we set

$$L_\theta f^\theta(x) = f_{(n)\theta}^\theta(x) + \sum_{r=0}^{n-1} p_r^\theta(x) f_{(r)\theta}^\theta(x),$$

$$W_\theta(x) = \sum_{r=0}^{n-1} |f_{(r)\theta}^\theta(x)|^2,$$

.5 will apply if only $C^\theta = \theta C$ is sufficiently small, which is indeed so for all $\theta < \epsilon$ for some $\epsilon > 0$ (remember C is fixed). Thus (II) gives for all $\theta < \epsilon$

$$\int_0^x dx_1 \int_0^{x_1} W_\theta(x_2)\,dx_2 \leqq \bar{K}\left\{\int_0^\infty |f^\theta|^2 + \int_0^\infty |L_\theta f^\theta|^2\right\}x + \bar{K}\theta C\int_0^\infty |f^\theta|^2 + \bar{K}\theta C\int_0^x W_\theta(x_1)\,dx_1 + \bar{K}W_\theta(x)$$

$$\leqq \bar{K}\left\{\int_0^\infty |f|^2 + \int_0^\infty |Lf|^2\right\}\frac{x}{\theta} + \bar{K}C\int_0^\infty |f|^2 + \bar{K}C\theta\int_0^x W_\theta(x_1)\,dx_1 + \bar{K}W_\theta(x)$$

since $\int_0^\infty |f^\theta|^2 = \frac{1}{\theta}\int_0^\infty |f|^2$ and $\int_0^\infty |L_\theta f^\theta|^2 = \int_0^\infty \theta^{2n-1}|Lf|^2 \leqq \frac{1}{\theta}\int_0^\infty |Lf|^2$.[23]

Now set $K_1 = \bar{K}\left\{\int_0^\infty |f|^2 + \int_0^\infty |Lf|^2\right\}$, $K_2 = \bar{K}C\int_0^\infty |f|^2$, $K_3 = \bar{K}C$, $K_4 = \bar{K}$, and define the function $\mathring{W}(x) = W_\theta(x/\theta)$. The previous inequality becomes, after x is replaced by x/θ,

$$(\mathrm{II}_\theta)\qquad \int_0^x dx_1 \int_0^{x_1} \mathring{W}(x_2)\,dx_2 \leqq K_1\theta x + K_2\theta^2 + K_3\theta^2\int_0^x \mathring{W}(x_1)\,dx_1 + K_4\theta^2\mathring{W}(x).$$

($\mathring{W}(x)$ depends on θ; K_1, K_2, K_3, K_4 are independent of θ.)

Since

$$\mathring{W}(x) = \sum_{r=0}^{n-1}\left|f^\theta_{(r)\theta}\left(\frac{x}{\theta}\right)\right|^2 = \sum_{r=0}^{n-1}\left|\theta^r f_{(r)}\left(\theta\frac{x}{\theta}\right)\right|^2 = \sum_{r=0}^{n-1}|\theta^r f_{(r)}(x)|^2 \geqq \theta^{2(n-1)}W(x),^{23}$$

$\int_0^\infty W(x)\,dx = \infty$ implies $\int_0^\infty \mathring{W}(x)\,dx = \infty$, that is, $\int_0^x \mathring{W}(x_1)\,dx_1 \to \infty$ as $x \to \infty$. Hence for every fixed $\theta(\theta < \epsilon)$ there will be a fixed $x_0(= x_0(\theta))$ such that $x \geqq x_0$ implies

$$K_1\theta x + K_2\theta^2 < \theta^2\int_0^x \mathring{W}(x_1)\,dx_1$$

and this, together with (II_θ), gives finally

$$(\mathrm{II}_{\theta x_0})\qquad \int_0^x dx_1\int_0^{x_1}\mathring{W}(x_2)\,dx_2 \leqq K_5^2\theta^2\left\{\int_0^x \mathring{W}(x_1)\,dx_1 + \mathring{W}(x)\right\}$$

for all $x \geqq x_0(\theta)$, for all $\theta < \epsilon$, where K_5 is a constant (independent of θ!).

.7 Consider the inequality ($\mathrm{II}_{\theta x_0}$) in .6; let θ be further restricted by the condition $\theta < 1/(2e)^{\frac{1}{2}}K_5$, and thus $1 > (2e)^{\frac{1}{2}}K_5\theta = s$ (say), and let the $x_0 = x_0(\theta)$ be chosen so large that $\int_0^{x_0} dx_1 \int_0^{x_1}\mathring{W}(x_2)\,dx_2 > 1$ (this is possible since we are assuming that $\int_0^\infty W(x)\,dx$, and with it $\int_0^\infty \mathring{W}(x)\,dx$, are infinite). We wish to prove the following set of propositions:

$$(P_0)\qquad \int_0^{x_0} dx_1\int_0^{x_1}\mathring{W}(x_2)\,dx_2 \geqq e^0 = 1$$

$$(P_m)\qquad \begin{cases} \int_0^x \mathring{W}(x_1)\,dx_1 \geqq \dfrac{e^m}{s} \quad \text{for } x_0 + (2m-1)s \leqq x \\ \int_0^{x_0+2ms} dx_1 \int_0^{x_1}\mathring{W}(x_2)\,dx_2 \geqq e^m \end{cases}$$

$$m = 1, 2, \cdots.$$

---

[23] $\theta < 1$!

Suppose (P_m) holds for some $m = 0, 1, \cdots$; then at least one of the following two statements must be true (use $(\mathrm{II}_{\theta x_0})$ and remember that $\int_0^x \mathring{W}(x_1)\, dx_1$, and also with it, $\int_0^x dx_1 \int_0^{x_1} \mathring{W}(x_2)\, dx_2$ are non-decreasing functions of x):

(i) $K_5^2 \theta^2 \mathring{W}(x) \geqq \frac{1}{2} e^m$ for $x_0 + 2ms \leqq x \leqq x_0 + (2m + 1)s$ or

(ii) $K_5^2 \theta^2 \int_0^x \mathring{W}(x_1)\, dx_1 \geqq \frac{1}{2} e^m$ for $x_0 + (2m + 1)s \leqq x$.

In either case, since $s < 1$,

$$\int_0^x \mathring{W}(x_1)\, dx_1 \geqq \frac{e^m s}{2K_5^2 \theta^2} = \frac{e^{m+1}}{s}$$

for $x_0 + (2m + 1)s \leqq x$, and hence

$$\int_0^{x_0+2(m+1)s} dx_1 \int_0^{x_1} \mathring{W}(x_2)\, dx_2 \geqq s\,\frac{e^{m+1}}{s} = e^{m+1}.$$

Thus (P_m) always implies (P_{m+1}), and (P_0) holds because of our choice of x_0. This means that (P_m) holds for all m.

Hence, for $x \geqq x_0 + s$, we have

$$\int_0^x \mathring{W}(x_1)\, dx_1 \geqq e^{\left[\frac{1}{2}\left(\frac{x-x_0}{s}+1\right)\right]} \qquad [t] = (\text{largest integer} \leqq t)$$

$$\geqq e^{-\left(\frac{x_0+s}{2}\right)} e^{\frac{x}{2s}}.$$

Since $W(x) \geqq \mathring{W}(x)$,[24] we can say that there is, for every fixed θ ($\theta < \epsilon$), an $\bar{x} = \bar{x}(\theta)$, such that $x \geqq \bar{x}$ implies

$$\int_0^x W(x_1)\, dx_1 \geqq F e^{\frac{x}{G\theta}} \tag{2}$$

where $F(>0)$ depends on θ, but G does not depend on θ! This is the second inequality mentioned in .3. It has been derived under the assumption that $\int_0^b W(x)\, dx = +\infty$, and it clearly contradicts the inequality (1) in .4 for sufficiently small θ. Thus $\int_0^b W(x)\, dx$ cannot be infinite and the lemma is established.

COROLLARY. *If the hypotheses of Lemma* 2.2 *hold on the interval* (a, ∞), *then* $f_{(r)}(x) \to 0$ *as* $x \to \infty$ *for* $0 \leqq r < n$.

---

[24] $\mathring{W}(x) = \sum_{r=0}^{n-1} |f^{\theta}_{(r)_\theta}(x/\theta)|^2 = \sum_{r=0}^{n-1} |\theta^r f_{(r)}(x)|^2$

$$\leqq \sum_{r=0}^{n-1} |f_{(r)}(x)|^2 = W(x), \quad \mathring{W}(x) \leqq W(x).$$

PROOF. For $0 \leqq r < n$, we have

$$|f_{(r)}(x_2)|^2 - |f_{(r)}(x_1)|^2 = \int_{x_1}^{x_2} \frac{d}{dx} |f_{(r)}(x)|^2 \, dx$$

$$= \int_{x_1}^{x_2} 2\Re \{\overline{f_{(r)}(x)} f'_{(r)}(x)\} \, dx = \int_{x_1}^{x_2} 2\Re \{\overline{f_{(r)}(x)} (f_{(r+1)}(x) - q_{r+1}(x)f(x))\} \, dx$$

$$\leqq \int_{x_1}^{x_2} 2 |f_{(r)}(x)| \, |f_{(r+1)}(x) - q_{r+1}(x)f(x)| \, dx$$

$$\leqq \int_{x_1}^{x_2} \{|f_{(r)}(x)|^2 + |f_{(r+1)}(x) - q_{r+1}(x)f(x)|^2\} \, dx$$

$$\leqq \int_{x_1}^{x_2} |f_{(r)}(x)|^2 \, dx + 2 \int_{x_1}^{x_2} |f_{(r+1)}(x)|^2 \, dx + 2C^2 \int_{x_1}^{x_2} |f(x)|^2 \, dx.$$

Now Lemma 2.2 implies that each term in the last expression converges to zero when x_1, x_2 tend to infinity. This means that $|f_{(r)}(x)|^2$ converges to a limit when x tends to infinity. But $\int_a^\infty |f_{(r)}(x)|^2 \, dx$ is finite, and so the limit cannot be different from zero. Hence $|f_{(r)}(x)|^2$, and with it $f_{(r)}(x)$, converge to zero when x tends to infinity.

REMARK. Let $b = 1$, and let H, D', $\bar{D}'$, D'_0, T', T'_0 be as in section 1, Definition 1.1′ (page 888), footnote (19) (page 890), and Definition 1.2′ (page 889), resp., but with $q_0(x)$, $p_n(x)$, $\equiv 1$. Then $T'f(x) = Lf(x)$ for every f in $\bar{D}'$. Suppose now that f_m, $m = 1, 2, \cdots$, is a sequence of elements in D' such that

$$\odot \qquad f_m \to f, \qquad T'f_m \to h,$$

hold simultaneously for some elements, f, h in H. We can write each $f_m(x)$ in the form $f_{m,1}(x) + f_{m,2}(x)$ where $f_{m,1}$ is in $\bar{D}'$, and $f_{m,2}$ is in D' with $T'f_{m,2} = 0$.[25] Then $Lf_{m,1}(x) = T'f_{m,1}(x) = T'f_m(x)$ will be a convergent

---

[25] Let P be the linear sub-space of the elements g in D' for which $T'g = 0$, and let $(f_m)_{(r)}(0) = a_r$. If $g_{(r)}(0) = a_r$ for some g in P, then clearly $f_m = (f_m - g) + g$ is a decomposition of the required kind. But such a g does exist, and it can be constructed (with some labor, but no difficulty) by applying the methods of successive substitutions to the system:
$g(x) = g_{(0)}(x)$

$$g_{(r)}(x) = a_r + \int_0^x (g_{(r+1)}(x_1) - q_{r+1}(x_1)g(x_1)) \, dx_1 \qquad \text{for } 0 \leqq r < n - 1.$$

$$g_{(n-1)}(x) = a_{n-1} + \int_0^x \left(-\sum_{r=0}^{n-1} p_r(x_1)g_{(r)}(x_1) - q_n(x_1)g(x_1)\right) dx_1.$$

Further, if g_1, g_2 were two solutions, then $g_1 - g_2$ would be in $\bar{D}'$, and $L(g_1(x) - g_2(x))$ would $= 0$. The inequality (1), page 894, then implies $g_1 = g_2$. Thus every g in P is uniquely characterized by its $g_{(r)}(0)$, $0 \leqq r < n$, and hence P is finite dimensional, with dimension $= n$.

sequence (in H). The inequality (1), page 894, together with the inequality (I), page 892, now show that $(f_{m,1})_{(n)}(x)$ converges (in H), and $(f_{m,1})_{(r)}(x)$ converges uniformly in x, for $0 \leqq r < n$, as m tends to infinity. From this follows that $f_{m,1}$ converges (in H) to a limit, say f_1, in $\bar{D}'$, such that $T'f_1 = h$.[26] Then $f_{m,2} = f_m - f_{m,1}$ converges in H to $f - f_1 = f_2$ (say). Now all $f_{m,2}$ lie in the linear sub-space of H formed by the solutions g of $T'g = 0$. This sub-space is finite dimensional,[25] hence closed, and thus contains f_2. This means, $T'f_2$ is defined and $= 0$, and finally, $T'f$ is defined and $= T'(f_1 + f_2) = T'f_1 = h$. Thus $\odot$ implies, $T'f$ is defined and $= h$. In other words, T' is closed.

If further all f_m are in D_0', then $f_m = f_{m,1}$ are all in D_0' and similar reasoning to that given above shows that the limit of the $f_{m,1}$ is also in D_0'. Hence T_0' is also closed.

The general case ($q_0(x)$, $p_n(x)$ not $\equiv 1$) can be reduced to the special case by replacing $f(x)$, $q_r(x)$, $p_r(x)$ by $q_0(x)f(x)$, $\dfrac{q_r(x)}{q_0(x)}$, $\dfrac{p_r(x)}{p_n(x)}$ resp. Hence T', T_0' are closed in the general case too.

For the remainder of this section H will denote the Hilbert space of functions L.s.s. on $(0, \infty)$ (any semi-infinite or infinite interval could be treated in the same way). Let $q_r(x)$, $p_r(x)$, D', D_0', T', T_0', D'', D_0'', T'', T_0'' be as in Theorem 1.1, page 000, but with $b = \infty$. We have then the theorem

THEOREM 1.2$'$. *Theorem* 1.1$'$ *remains valid.*

PROOF. If $(T_0')^*g = g^*$, then $\int_0^\infty (T_0'h\cdot\bar{g} - h\cdot\overline{g^*}) = 0$ for all h in D_0'. Now any $f(x)$ in the D_0' of Theorem 1.1 can be extended to an $h(x)$ in the D_0' of the present theorem by defining $h(x) = f(x)$ for $0 \leqq x \leqq 1$, and $h(x) = 0$ for $x > 1$. Then $\int_0^1 (T_0'f\cdot\bar{g} - f\cdot\overline{g^*}) = \int_0^\infty (T_0'h\cdot\bar{g} - h\cdot\overline{g^*}) = 0$ for all such f. Theorem 1.1$'$ now gives that $g^*(x) \sim \sum_{r=0}^n (-1)^{n-r}\overline{q_{n-r}(x)}\,g_{(r)}(x)$ for $0 \leqq x \leqq 1$. Since this equivalence must hold, in the same way, for every finite interval $(0, b)$, it holds even on the semi-infinite interval $(0, \infty)$. Further, g and $g^* = \sum_{r=0}^n (-1)^{n-1}\overline{q_{n-r}}g_{(r)}$ are both L.s.s. on $(0, \infty)$, which, by Lemma 2.2, implies

---

[26] Let the uniform limit of $(f_{m,1})_{(r)}(x)$ be denoted by $k_r(x)$, $0 \leqq r < n$, and let the limit (in H) of $(f_{m,1})_{(n)}$ be denoted by k_n. Then clearly the $k_r(x)$ satisfy the equations

$$k_r(0) = 0, \qquad k_r(x) = \int_0^x (k_{r+1}(x_1) - q_{r+1}(x_1)k_0(x_1))\,dx_1 \qquad 0 \leqq r < n$$

since the $(f_{m,1})_{(r)}(x)$ do so for every m. But these equations imply

$$k_r(0) = 0, \qquad k_{r+1}(x) \sim k_r(x) + q_{r+1}(x)k_0(x) \qquad 0 \leqq r < n$$

that is, $k(x) = k_0(x)$ is in $\bar{D}'$, and $k_{(r)}(x) = k_r(x)$. Then $(f_{m,1})_{(r)}(x)$ converges uniformly in x to $k_{(r)}(x)$ for $0 \leqq r < n$, and $(f_{m,1})_{(n)}$ converges (in H) to $k_{(n)}$. Since the $p_r(x)$ are bounded, this implies that $Lf_{m,1}$ converges (in H) to Lk. But $Lf_{m,1}$ converges to h. Hence $Lk = h$. Writing f_1 in place of k, we obtain the desired statement.

that g must be in the D'' of the present theorem. Thus $(T_0')^* \subset T''$. Similarly $(T')^* \subset T_0''$. But the identity

$$\int_0^\infty (T'h\cdot\bar{g} - h\cdot\overline{T''g}) = \sum_{r=0}^{n-1} (-1)^{r+1}(f_{(n-r-1)}(0)\overline{g_{(r)}(0)})$$

for all h in D' and all g in D'',[27] shows that $(T')^* \supset T_0''$, $(T_0')^* \supset T''$. Hence $(T')^* = T_0''$, $(T_0')^* = T''$ and the theorem follows.[28]

3. The Hyperbolic Differential Operator

In this § we study the expression

$$A(x, y)\frac{\partial^2}{\partial x\,\partial y} + B(x, y)\frac{\partial}{\partial x} + C(x, y)\frac{\partial}{\partial y} + E(x, y).$$

Similar methods and results hold for the general

$$p_{n_1,n_2,\cdots,n_t}(x_1, x_2, \cdots, x_t)\frac{\partial^{n_1+n_2+\cdots+n_t}}{\partial x_1^{n_1}\,\partial x_2^{n_2}\cdots\partial x_t^{n_t}}$$
$$+ \sum_{r_k<n_k} p_{r_1,r_2,\cdots,r_t}(x_1, x_2, \cdots, x_t)\frac{\partial^{r_1+r_2+\cdots+r_t}}{\partial x_1^{r_1}\,\partial x_2^{r_2}\cdots\partial x_t^{r_t}}.$$

H will denote the Hilbert space of functions $f(x, y)$ which are L.s.s. on the square I: $0 \leqq x, y \leqq 1$. We shall assume that A, B, C, E are measurable and bounded on I and that $|A(x, y)| \geqq \epsilon > 0$.

Define sub-spaces $D^{(3)}$, $D_0^{(3)}$ in H by

Definition 3.1. $D^{(3)}$ is the linear space of the $f(x, y)$ which have the form

$$f(x, y) = f(0, 0) + \int_0^x \varphi_1(w)\,dw + \int_0^y \varphi_2(z)\,dz + \int_0^x\int_0^y \varphi_3(w, z)\,dz\,dw$$

where $f(0, 0)$ is a constant and φ_1, φ_2, φ_3 are L.s.s. on I.

Definition 3.2. $D_0^{(3)}$ is the linear space of the $f(x, y)$ in $D^{(3)}$ which vanish on the boundary of I.

Define operators $T^{(3)}$, $T_0^{(3)}$ with domains $D^{(3)}$, $D_0^{(3)}$ resp. by $T_0^{(3)} \subset T^{(3)}$ and

$$T^{(3)}f(x, y) = A(x, y)f_{xy}(x, y) + B(x, y)f_x(x, y) + C(x, y)f_y(x, y) + E(x, y)f(x, y).$$

The function $T^{(3)}f(x, y)$ will clearly be defined almost everywhere and will be L.s.s.[29]

---

[27] Use integration by parts and apply the Corollary to Lemma 2.2.

[28] This theorem, for finite and infinite intervals, as well as the other results of sections 1 and 2, can be proved with the same methods, for the general quasi-differential operator where the $f_{(r)}(x)$ are required to satisfy equivalences $f_{(r+1)}(x) \sim f_{(r)}'(x) + \sum_{R=0}^{n} q_{r,R}(x)f_{(R)}(x)$.

[29] If $h(x) \sim k(x)$ with $k(x)$ abs. cont., then we define $h'(x)$ as any function $\sim k'(x)$.

THEOREM 3.1. *$T_0^{(3)}$ is a closed operator. The closure of $T^{(3)}$ is defined for exactly those functions $f(x, y)$ which have the form $f_1(x, y) + f_2(x, y)$ where f_1 is the limit of a sequence of elements in $D^{(3)}$ for which $T^{(3)}$ vanishes, and f_2 is in $D^{(3)}$; the value of the closure will be given by $\tilde{T}^{(3)}f = T^{(3)}f_2$.*

For the proof of this theorem we require the

LEMMA 3.1. *If* (i) $f(x, y) = \int_0^x \int_0^y f_{wz}(w, z)\, dz\, dw$ *where f_{xy} is L.s.s. on every* $0 \leqq x \leqq a < 1; 0 \leqq y \leqq b < 1$, *and*

(ii) $Lf(x, y) = A(x, y)f_{xy}(x, y) + B(x, y)f_x(x, y) + C(x, y)f_y(x, y) + E(x, y)f(x, y)$

is L.s.s. on $0 \leqq x, y \leqq 1$, where $|B|, |C|, |E|$, are less than some finite K, and $|A| \geqq \epsilon > 0$ on $0 \leqq x, y \leqq 1$, then f_{xy} is L.s.s. on $0 \leqq x, y \leqq 1$ and

$$\int_0^1 \int_0^1 |f_{wz}(w, z)|^2 dz\,dw \leqq K_1 \int_0^1 \int_0^1 |Lf(w, z)|^2 dz\,dw$$

where K_1 is a constant depending only on K.

PROOF. Let $0 \leqq x, y < 1$ and let

$$\theta(x, y) = 4 \int_0^x \int_0^y \{|Lf|^2 + K^2|f_x|^2 + K^2|f_y|^2\}.$$

Then

$$\int_0^x \int_0^y |f_{xy}|^2 = \int_0^x \int_0^y |Lf - Bf_x - Cf_y - Ef|^2$$

$$\leqq 4 \int_0^x \int_0^y \{|Lf|^2 + K^2|f_x|^2 + K^2|f_y|^2 + K^2|f|^2\}$$

$$\leqq \theta(x, y) + 4K^2 \int_0^x \int_0^y \left| \int_0^{x_1} \int_0^{y_1} f_{xy} \right|^2 \leqq \theta(x, y) + 4K^2 xy \int_0^x \int_0^y \cdot \int_0^{x_1} \int_0^{y_1} |f_{xy}|^2.$$

Iterating this inequality and noting that

$$(4K^2xy)^N \int \int_0^y \int_0^{x_1} \int_0^{y_1} \cdots \int_0^{x_N} \int_0^{y_N} |f_{xy}|^2 \leqq (4K^2xy)^N \int_0^x \int_0^y |f_{xy}|^2 \frac{(xy)^N}{N!\,N!} \to 0$$

as $N \to \infty$, we obtain

$$\int_0^x \int_0^y |f_{xy}|^2 \leqq K_2\theta(x, y)$$

where K_2 is a constant (depending only on K).[30]

[30] The symbol K with a subscript will be used to denote (during this proof) some suitable constant, which depends only on K.

Now let

$$W_1(x, y) = \int_0^x |f_x|^2, \qquad W_2(x, y) = \int_0^y |f_y|^2.$$

From

$$|(|f_x|^2)_y| = |\{f_{xy}\bar{f}_x + \overline{f_{xy}}f_x\}| \leqq |f_{xy}|^2 + |f_x|^2$$

we get by a double integration

$$\int_0^x |f_x|^2 \leqq \int_0^x \int_0^y |f_{xy}|^2 + \int_0^x \int_0^y |f_x|^2 \leqq M + K_3 \int_0^x \int_0^y |f_y|^2 + K_3 \int_0^x \int_0^y |f_x|^2,$$

where

$$M = 4K_2 \int_0^1 \int_0^1 |Lf|^2, \qquad K_3 = 1 + 4K_2K^2.$$

Hence

$$W_1(x, y) - K_3 \int_0^y W_1(x, z)\, dz \leqq M + K_3 \int_0^x W_2(w, y)\, dw.$$

Similarly,

$$W_2(x, y) - K_3 \int_0^x W_2(w, y)\, dw \leqq M + K_3 \int_0^y W_1(x, z)\, dz.$$

From these inequalities we obtain

$$\frac{\partial}{\partial y}\left\{e^{-K_3y} \int_0^y W_1(x, z)\, dz\right\} \leqq e^{-K_3y}\left\{M + K_3 \int_0^x W_2(w, y)\, dw\right\},$$

$$\int_0^y W_1(x, z)\, dz \leqq e^{K_3y}\left\{\frac{M}{K^3}(1 - e^{-K_3y}) + K_3 \int_0^x \int_0^y W_2(w, z)\, dz\, dw\right\}$$

$$\leqq K_4M + K_4 \int_0^x \int_0^y W_2(w, z)\, dz\, dw,$$

and similarly,

$$\int_0^x W_2(w, y)\, dw \leqq K_4M + K_4 \int_0^x \int_0^y W_1(w, z)\, dz\, dw.$$

If we set $S(x, y) = \int_0^y W_1 = \int_0^x \int_0^y |f_x|^2$, then the last two inequalities give

$$S(x, y) \leqq K_4M + K_4 \int_0^y \left\{K_4M + K_4 \int_0^x \int_0^{y_1} W_1\right\}$$

$$\leqq K_5M + K_5 \int_0^x \int_0^y S.$$

Iterating this inequality,

$$S(x, y) \leqq K_5M + K_5\int_0^x\int_0^y K_5M + K_5^2\int_0^x\int_0^y\int_0^{x_1}\int_0^{y_1} S \leqq \cdots$$

$$\leqq K_5M + K_5\int_0^x\int_0^y K_5M + \cdots$$

$$+ K_5^{N+1}\int_0^x\int_0^y\int_0^{x_1}\int_0^{y_1}\cdots\int_0^{x_N}\int_0^{y_N} K_5M + \cdots$$

$$\leqq K_5M\left\{1 + K_5xy + \cdots + K_5^{N+1}\frac{x^{N+1}y^{N+1}}{(N+1)!\,(N+1)!} + \cdots\right\}$$

$$\leqq K_5Me^{K_5xy} \leqq K_6M,$$

$$\int_0^x\int_0^y |f_x|^2 \leqq K_6M.$$

Similarly, $\int_0^x\int_0^y |f_y|^2 \leqq K_6M$. Finally,

$$\int_0^1\int_0^1 |f_{xy}|^2 \leqq 4K_2\int_0^1\int_0^1 \{|Lf|^2 + K^2|f_x|^2 + K^2|f_y|^2\} \leqq K_1M$$

$$= K_1\int_0^1\int_0^1 |Lf|^2$$

and the proof of the lemma is completed.

Proof of Theorem 3.1. If f_m is a sequence of elements from $D_0^{(3)}$ such that

$$f_m \to f \qquad\qquad T_0^{(3)}f_m \to g$$

hold simultaneously when m tends to infinity, for some f, g in H, then Lemma 3.1 shows that $(f_m)_x$, $(f_m)_y$, $(f_m)_{xy}$, converge in H. Hence $f_m(x, y)$ converges uniformly in x, y to an $f(x, y)$ in $D_0^{(3)}$, and $g = T_0^{(3)}f$, proving that $T_0^{(3)}$ is closed.

If f_m is a sequence of elements from $D^{(3)}$, each f_m can be put in the form $f_m = f_{m,1} + f_{m,2}$ where $f_{m,1}$ is an element of $D^{(3)}$ such that $T^{(3)}f_{m,1} = 0$, and $f_{m,2}(x, y)$ is in $D^{(3)}$ and vanishes when $x = 0$ or $y = 0$ and hence is of the form to which Lemma 3.1 applies.[31] Now if the f_m are such that $f_m \to f$ and $T^{(3)}f_m \to g$ for some f, g in H, as m tends to infinity, then $T^{(3)}f_{m,2} = T^{(3)}f_m$ converges to g. Lemma 3.1 now gives that $f_{m,2}$ converges to some f_2 in $D^{(3)}$ such that $T^{(3)}f_2 = g$, and, consequently, $f_{m,1}$ converges also (to $f - f_2$). This completes the proof of the theorem.[32]

---

[31] A well known existence theorem in differential equations (see J. Horn: *Partielle Differentialgleichungen*, 1929, p. 16) shows that there is an $f_{m,2}(x, y)$ in $D^{(3)}$ such that $T^{(3)}f_{m,2} = T^{(3)}f_m$ and $f_{m,2}(x, y) = 0$ if $x = 0$ or $y = 0$. The required decomposition will then be obtained by setting $f_{m,1} = f_m - f_{m,2}$.

[32] Let M be the linear subspace of $D^{(3)}$ consisting of the elements f for which $T^{(3)}f = 0$. Then Theorem 3.1 says that the domain of $T^{(3)}$ is the smallest linear space which contains $D^{(3)}$ and all limit points of M. The differential expression may not apply directly to the functions determining these limit elements of M, as can be shown by simple examples (see Theorem 3.2, and footnote (36) for a deeper discussion).

Let g, g^* be L.s.s. If we write every f in $D^{(3)}$ in the form given in Definition 3.1, the method of Theorem 1.1 shows that $\int\int_I T^{(3)}f\cdot\bar{g} = \int\int_I f\cdot\overline{g^*}$ for every f in $D^{(3)}$, is equivalent to

$$\overline{A(x,y)}\,g(x,y) + \int_y^1 \overline{B(x,z)}g(x,z)\,dz + \int_x^1 \overline{C(w,y)}g(w,y)\,dw + \int_x^1\int_y^1 \overline{E(w,z)}g(w,z)\,dz\,dw \sim \int_x^1\int_y^1 g^*(w,z)\,dz\,dw \tag{1}$$

$$\int_0^1 \overline{C(w,y)}g(w,y)\,dw + \int_y^1\int_0^1 \overline{E(w,z)}g(w,z)\,dz\,dw \sim \int_y^1\int_0^1 g^*(w,z)\,dz\,dw \qquad (\sim \text{ in } y!) \tag{2}$$

$$\int_0^1 \overline{B(x,z)}g(x,z)\,dz + \int_x^1\int_0^1 \overline{E(w,z)}g(w,z)\,dz\,dw \sim \int_x^1\int_0^1 g^*(w,z)\,dz\,dw \qquad (\sim \text{ in } x!) \tag{3}$$

$$\int_0^1\int_0^1 \overline{E(w,z)}g(w,z)\,dz\,dw = \int_0^1\int_0^1 g^*(w,z)\,dz\,dw. \tag{4}$$

To avoid the complications of quasi-integro-differential operators we shall impose further restrictions on the coefficients A, B, C, E. Call $\chi(x, y)$ *regularly absolutely continuous with respect to* x (abbrev. reg. abs. cont. w.r.t. x) in a region S, if for every $\epsilon > 0$ there is a $\rho(\epsilon) > 0$ such that

$$\textstyle\sum_{m=1}^{\infty} |\chi(x_m, y_m) - \chi(x_{m'}, y_m)| < \epsilon$$

whenever $\sum_{m=1}^{\infty} |x_m - x_{m'}| < \rho(\epsilon)$ and all (x_m, y_m), $(x_{m'}, y_m)$ are in S. We shall assume that

(i) A, A_y, B are defined almost everywhere in I, are measurable and bounded and reg. abs. cont. w.r.t. x.

(ii) A, A_x, C are defined almost everywhere in I, are measurable and bounded and reg. abs. cont. w.r.t. y.

(iii) A_{xy}, C_y, B_x, E, are defined almost everywhere in I and are measurable and bounded.

With these assumptions we can define operators $T^{(3')}$, $T_0^{(3')}$ on the domains $D^{(3)}$, $D_0^{(3)}$ respectively, by $T_0^{(3')} \subset T^{(3')}$ and

$$T^{(3')}f = A^*f_{xy} + B^*f_x + C^*f_y + E^*f \equiv (\bar{A}f)_{xy} - (\bar{B}f)_x - (\bar{C}f)_y + (\bar{E}f).$$

THEOREM 3.2. *$T_0^{(3')}$ is the adjoint of $T^{(3)}$ and $\bar{T}^{(3')}$ is the adjoint of $T_0^{(3)}$. An element g is in the domain of $\bar{T}^{(3')}$ if and only if g has the form $g_1 + g_2$, where*

$$\overline{A(x,y)}\,g_1(x,y) + \int_y^1 \overline{B(x,z)}g(x,z)\,dz + \int_x^1 \overline{C(w,y)}g(w,y)\,dw + \int_x^1\int_y^1 \overline{E(w,z)}g(w,z)\,dz\,dw \sim h_1(x) + h_2(y) \tag{i}$$

for L.s.s. functions $h_1(x)$, $h_2(y)$.

(ii) $g_2(x, y) = \int_x^1 \int_y^1 \varphi(w, z)\, dz\, dw$ *with* $\varphi(=(g_2)_{xy})$ *L.s.s.* (g_2 *will be in* $D^{(3)}$).

Every set of L.s.s. functions $\varphi(x, y)$, $h_1(x)$, $h_2(y)$, *will determine uniquely a* g *in the domain of* $\tilde{T}^{(3')}$ *and* $\tilde{T}^{(3')}g$ *will equal* $T^{(3')}g_2$.[33]

For the proof of Theorem 3.2 we require the following lemmas which are based on the method of successive substitutions. We prove these lemmas for any region S which is finite and bounded by curves:

$$\begin{array}{llll} y = c & a \leqq x \leqq b; & y = \delta(x) & a \leqq x \leqq b; \\ x = a & c \leqq y \leqq d; & x = b & c \leqq y \leqq \delta(b); \end{array}$$

or, equivalently, by curves:

$$\begin{array}{llll} x = a & c \leqq y \leqq d; & x = \beta(y) & c \leqq y \leqq d; \\ y = c & a \leqq x \leqq b; & y = d & a \leqq x \leqq \beta(d); \end{array}$$

where $\delta(x)$, $\beta(y)$ are non-increasing, with $d = \delta(a) \geqq \delta(b) \geqq c$ and $b = \beta(c) \geqq \beta(d) \geqq a$.

LEMMA 3.2. *If* $A(x, y)\cdot g(x, y) \sim \int_y^{\delta(x)} C(x, z)g(x, z)\, dz + \int_x^{\beta(y)} \varphi(w, y)\, dw$ *where*

(i) $|A(x, y)| \geqq \epsilon > 0$,
(ii) Ag, C, φ *are summable on* S,
(iii) C/A *is reg. abs. cont. w.r.t.* x *and less than* K *in absolute value*,

then Ag, *or an equivalent function, is abs. cont. in* x *for fixed* y *and has a partial derivative w.r.t.* x *which is summable over* S.

PROOF. We may clearly assume $A(x, y) \equiv 1$. The formulae

$$\lambda_m(x, y) = \int_y^{\delta(x)} C(x, z_1)\, dz_1 \int_{z_1}^{\delta(x)} \cdots \int_{z_{m-1}}^{\delta(x)} C(x, z_m)\, dz_m \int_x^{\beta(z_m)} \varphi(w, z_m)\, dw$$

$$\mu_m(x, y) = \int_y^{\delta(x)} C(x, z_1)\, dz_1 \int_{z_1}^{\delta(x)} \cdots \int_{z_{m-1}}^{\delta(x)} C(x, z_m)\, dz_m \int_{z_m}^{\delta(x)} C(x, z_{m+1})g(x, z_{m+1})\, dz_{m+1}$$

define $\lambda_m(x, y)$ everywhere and $\mu_m(x, y)$ almost everywhere in S and

$$g(x, y) \sim \int_x^{\beta(y)} \varphi(w, y)\, dw + \sum_{m=1}^{N} \lambda_m(x, y) + \mu_N(x, y) \qquad N = 1, 2, \cdots.$$

Well known methods of evaluation give

$$|\mu_N(x, y)| \leqq K^{N+1} \frac{(\delta(x) - y)^N}{N!} \int_a^{(\delta)x} |g(x, z)|\, dz$$

---

[33] The corresponding statement about the domain of $\tilde{T}^{(3)}$ can be easily obtained by simply interchanging the rôles of $T^{(3)}$ and $T^{(3')}$.

which converges to zero almost everywhere in S as N tends to infinity, and

$$|\lambda_m(x, y)| \leqq K^m \frac{(\delta(x) - y)^{m-1}}{(m-1)!} \int\int |\varphi(w, z)| \, dz \, dw,$$

implying $\lambda(x, y) = \sum_{m=1}^{\infty} \lambda_m(x, y)$ is uniformly and absolutely convergent in S. Thus

$$g(x, y) \sim \int_x^{\beta(y)} \varphi(w, y) \, dw + \lambda(x, y),$$

and we need only show that $\lambda(x, y)$ is abs. cont. in x and has a partial derivative w.r.t. x summable over S.

If x_j is the smaller of x_i, x_j and $c < y < \delta(x_j)$ we have

$$\begin{aligned}
&|\lambda_m(x_j, y) - \lambda_m(x_i, y)| \\
&\leqq \left| \int^{\delta(x_j)} C(x_j, z_1) \, dz_1 \int_{z_1}^{\delta(x_j)} \cdots \int_{z_{m-1}}^{\delta(x_j)} C(x_j, z_m) \, dz_m \int_{x_j}^{\beta(z_m)} \varphi(w, z_m) \, dw \right. \\
&\qquad \left. - \int_y^{\delta(x_i)} C(x_j, z_1) \, dz_1 \int_{z_1}^{\delta(x_i)} \cdots \int_{z_{m-1}}^{\delta(x_i)} C(x_j, z_m) \, dz_m \int_{x_i}^{\beta(z_m)} \varphi(w, z_m) \, dw \right| \\
&\qquad + \left| \int_y^{\delta(x_i)} C(x_j, z_1) \, dz_1 \int_{z_1}^{\delta(x_i)} \cdots \int_{z_{m-1}}^{\delta(x_i)} C(x_j, z_m) \, dz_m \int_{x_i}^{\beta(z_m)} \varphi(w, z_m) \, dw \right. \\
&\qquad \left. - \int_y^{\delta(x_i)} C(x_i, z_1) \, dz_1 \int_{z_1}^{\delta(x_i)} \cdots \int_{z_{m-1}}^{\delta(x_i)} C(x_i, z_m) \, dz_m \int_{x_i}^{\beta(z_m)} \varphi(w, z_m) \, dw \right|.
\end{aligned}$$

The first term on the right side of the last inequality is majorized by

$$\begin{aligned}
&\int_{\delta(x_i)}^{\delta(x_j)} K dz_1 \int_{z_1}^{\delta(x_j)} K dz_2 \cdots \int_{x_j}^{\beta(z_m)} |\varphi(w, z_m)| \, dw \\
&\qquad + \int_y^{\delta(x_i)} K dz_1 \int_{\delta(x_i)}^{\delta(x_j)} K dz_2 \int_{z_2}^{\delta(x_j)} \cdots \int_{x_j}^{\beta(z_m)} |\varphi(w, z_m)| \, dw + \cdots \\
&\qquad + \int_y^{\delta(x_j)} K dz_1 \int_{z_1}^{\delta(x_i)} K dz_2 \cdots \int_{z_{m-1}}^{\delta(x_i)} K dz_m \int_{x_j}^{x_i} |\varphi(w, z_m)| \, dw,
\end{aligned}$$

and so also by

$$\begin{aligned}
&K^m \varphi \left\{ \frac{\Delta^{m-1}}{(m-1)!} + \frac{l}{1!} \frac{\Delta^{m-2}}{(m-2)!} + \cdots + \frac{l^{m-2}}{(m-2)!} \cdot \frac{\Delta}{1!} \right\} \\
&\qquad + \frac{K^m l^{m-1}}{(m-1)!} \int_{x_j}^{x_i} dw \int_c^{\delta(w)} |\varphi(w, z)| \, dz,
\end{aligned}$$

where[34]

$$\varphi = \int\int_S |\varphi(x, y)| \, dy \, dx, \qquad l = d - c, \qquad \Delta = x_i - x_j.$$

---

[34] The subscript is ordinarily to be taken as indicating partial differentiation but there will be no confusion here.

In the second term we replace $C(x_j, z)$ by $C(x_i, z) + \theta(z)$ and set

$$\theta = \overline{\text{bd}}.\,|\,\theta(z)\,|, \qquad \eta_1(y) = \theta\cdot\varphi,$$

$$\eta_m(y) = (\theta + K)\int_y^{\delta(x_i)} \eta_{m-1}(z)\,dz + \theta\cdot K^{m-1}\cdot\varphi\cdot\frac{(\delta(x) - y)^{m-1}}{(m-1)!}, \quad (m > 1).$$

Then

$$|\,\lambda(x_j, y) - \lambda(x_i, y)\,| \leqq \sum_{m=1}^{\infty} |\,\lambda_m(x_j, y) - \lambda_m(x_i, y)\,|$$

$$\leqq K^2\cdot\varphi\cdot\Delta e^{K\Delta}e^{Kl} + \sum_{m=1}^{\infty}\eta_m(y) + Ke^{Kl}\int_{x_j}^{x_i} dw \int_C^{\delta(W)} |\,\varphi(w, z)\,|\,dz.$$

The inequality

$$(\eta) \qquad \eta_m(y) \leqq \theta\cdot\varphi\cdot m(3k)^{m-1}\cdot\frac{(\delta(x_i) - y)^{m-1}}{(m-1)!}$$

clearly holds for $m = 1$. If it holds for m, then it will hold for $m + 1$ also since

$$\eta_{m+1}(y) \leqq \{3K\theta\varphi m(3K)^{m-1} + \theta K^m\varphi\}\,\frac{(\delta(x_i) - y)^{m-1}}{(m-1)!}$$

$$\leqq \theta\cdot\varphi(m+1)(3K)^m\,\frac{(\delta(x_i) - y)^m}{m!}.$$

Thus (η) holds for all m and $\sum_{m=1}^{\infty}\eta_m(y) \leqq \theta\cdot M_3$ where M_3 is a constant. Defining the abs. cont., additive, interval-function $V(s, t)$ by

$$V(s,t) = \overline{\text{bd}}_{N<\infty}\left\{\sum_{m=1}^{N}\left[\overline{\text{bd}}_{z}\,|\,C(w_{m+1}, z) - C(w_m, z)\,|\,\right], s = w_1 < w_2 < \cdots < w_{N+1} = t\right\},$$

we obtain

$$|\,\lambda(x_j, y) - \lambda(x_i, y)\,| \leqq M_1\Delta + M_2\int_{x_j}^{x_i}\int_c^d |\,\varphi(w, z)\,|\,dz\,dw + M_3V(x_j, x_i)$$

where M_1, M_2, M_3 are independent of x_i, x_j and y. The last inequality shows that $\lambda(x, y)$ is abs. cont. in x and has a partial derivative with respect to x which is summable over S.[35]

[35] If $k(x_1, x_2, \cdots, x_t)$is measurable as a function of all the variables and continuous in x_1 for fixed values of the other variables, then the partial derivative k_{x_1} as far as it exists, is always measurable as a function of all the variables (see C. Carathéodory: *Reelle Funktionen*, 1918, p. 642, for a proof). Thus λ_x is measurable in x, y.

LEMMA 3.3. *If*

$$A(x, y)g(x, y) \sim B_1(x, y) \int_a^x B_2(w, y)g(w, y)\, dw + C_1(x, y) \int_c^y C_2(x, z)g(x, z)\, dz + \varphi(x, y)$$

on S, where A is measurable and $| A(x, y) | \geqq \epsilon > 0$; B_1, B_2, C_1, C_2 *are measurable and bounded; g is summable; and* φ *is L.s.s.; then g is L.s.s.*

PROOF. We may assume $A(x, y) \equiv 1$, and $| B_1 |, | B_2 |, | C_1 |, | C_2 |, \int\int_S | g |$, less than K. Then

$$g(x, y) \sim \varphi(x, y) + \sum_{m=1}^{N} \lambda_m(x, y) + \mu_N(x, y), \qquad N = 1, 2, \cdots,$$

where

$$\begin{aligned}\lambda_m(x, y) = {} & B_1(x, y) \int_a^x B_2(w_1, y) \cdot B_1(w_1, y)\, dw_1 \int_a^{w_1} \cdots \\ & \int_a^{w_{m-2}} B_2(w_{m-1}, y) \cdot B_1(w_{m-1}, y)\, dw_{m-1} \int_a^{w_{m-1}} B_2(w_m, y)\varphi(w_m, y)\, dw_m \\ & + B_1(x, y) \int_a^x B_2(w_1, y) \cdot B_1(w_1, y)\, dw_1 \int_a^{w_1} \cdots \\ & \int_a^{w_{m-2}} B_2(w_{m-1}, y)C_1(w_{m-1}, y)\, dw_{m-1} \int_c^y C_2(w_{m-1}, z)g(w_{m-1}, z)\, dz \\ & + C_1(x, y) \int_c^y C_2(x, z_1) \cdot C_1(x, z_1)\, dz_1 \int_c^{z_1} \cdots \\ & \int_c^{z_{m-2}} C_2(x, z_{m-1}) \cdot C_1(x, z_{m-1})\, dz_{m-1} \int_c^{z_{m-1}} C_2(x, z_m)\varphi(x, z_m)\, dz_m \\ & + C_1(x, y) \int_c^y C_2(x, z_1)C_1(x, z_1)\, dz_1 \int_c^{z_1} \cdots \\ & \int_c^{z_{m-2}} C_2(x, z_{m-1}) \cdot B_1(x, z_{m-1})\, dz_{m-1} \int_c^x B_2(w, z_{m-1})g(w, z_{m-1})\, dw.\end{aligned}$$

$$\begin{aligned}\mu_m(x, y) = {} & B_1(x, y) \int_a^x B_2(w_1, y) \cdot B_1(w_1, y)\, dw_1 \int_a^{w_1} \cdots \\ & \int_a^{w_{m-1}} B_2(w_m, y) \cdot B_1(w_m, y)\, dw_m \int_a^{w_m} B_2(w_{m+1}, y) \cdot g(w_{m+1}, y)\, dw_{m+1} \\ & + C_1(x, y) \int_c^y C_2(x, z_1) \cdot C_1(x, z_1)\, dz_1 \int_c^{z_1} \cdots \\ & \int_c^{z_{m-1}} C_2(x, z_m) \cdot C_1(x, z_m)\, dz_m \int_c^{z_m} C_2(x, z_{m+1})g(x, z_{m+1})\, dz_{m+1}.\end{aligned}$$

Familiar evaluations show that $\mu_N(x, y)$ converges to zero almost everywhere as N tends to infinity, and

$$g(x, y) - \varphi(x, y) \sim \sum_{i=1}^{4} \sum_{m=1}^{\infty} \lambda_m^{(i)}(x, y)$$

where $\lambda_m^{(i)}(x, y)$ is the i^{th} term in the expression defining $\lambda_m(x, y)$ above. We shall show that $\sum_{m=1}^{\infty} | \lambda_m^{(i)}(x, y) |$ is convergent almost everywhere in S with a L.s.s. sum for $i = 1, 2, 3, 4$. Due to the symmetry we need consider only $i = 1$ and $i = 2$. Since $| \lambda_m^{(1)}(x, y) | \leqq K^{2m} \dfrac{(x - a)^{m-1}}{(m - 1)!} \displaystyle\int_a^{\beta(y)} | \varphi(w, y) | \, dw$,

$$\left| \sum_{m=1}^{\infty} \lambda_m^{(1)}(x, y) \right| \leqq K^2 e^{K^2(\beta(y) - a)} \int_a^{\beta(y)} | \varphi(w, y) | \, dw$$

is finite almost everywhere in S, and

$$\int \int_S \left| \sum_{m=1}^{\infty} \lambda_m^{(1)}(x, y) \right|^2 dy\,dx \leqq M \int \int_S | \varphi(x, y) |^2 \, dy\,dx$$

where M is a constant. Hence $\sum_{m=1}^{\infty} | \lambda_m^{(1)}(x, y) |$ converges almost everywhere in S and has a L.s.s. sum. In the same way we derive from $| \lambda_m^{(2)}(x, y) | \leqq K^{2m+1} (x - a)^{m-1}/(m - 1)!$ that $\sum_{m=1}^{\infty} | \lambda_m^{(2)}(x, y) |$ converges everywhere in S and has a L.s.s. sum.

Thus in the equivalence

$$g(x, y) \sim \varphi(x, y) + \sum_{i=1}^{4} \lambda^{(i)}(x, y)$$

each term on the left is L.s.s. and so $g(x, y)$ is L.s.s.

Lemma 3.4. *If A is measurable and $| A(x, y) | \geqq \epsilon > 0$, and B, C are measurable and bounded, and E, φ are L.s.s.; then the equivalence*

$$A(x, y)f(x, y) \sim \varphi(x, y) + \int_a^x C(w, y)f(w, y)\,dw + \int_c^y B(x, z)f(x, z)\,dz + \int_a^x \int_c^y E(w, z)f(w, z)\,dz\,dw$$

has a L.s.s. solution $f(x, y)$, unique to within an equivalence in the sense of Lebesgue.

Proof. We may assume $A \equiv 1$. Define

$$f_0(x, y) = \varphi(x, y)$$

$$f_m(x, y) = \int_a^x C(w, y)f_{m-1}(w, y)\,dw + \int_c^y B(x, z)f_{m-1}(x, z)\,dz + \int_a^x \int_c^y E(w, z)f_{m-1}(w, z)\,dz\,dw$$

for $m = 1, 2, \cdots$, and let $f(x, y) = \sum_{m=0}^{\infty} f_m(x, y)$. Let

$$\varphi_x = \int_c^{\delta(x)} | \varphi(x, z) | \, dz, \qquad \varphi_y = \int_a^{\beta(y)} | \varphi(w, y) | \, dw,$$ [34]

$$\varphi = \max.\left\{\int \int_S | \varphi(w, z) | \, dz\, dw, \left(\int \int_S | \varphi(w, z) |^2 dz\, dw\right)^{\frac{1}{2}}\right\},$$

$$K = \overline{\mathrm{bd}}_{S}\left\{| B(x, y) |, | C(x, y), \left(\int \int_S | E(w, z) |^2 dz\, dw\right)^{\frac{1}{2}}\right\}.$$

When $f_m(x, y)$ is expressed explicitly in terms of B, C, E and φ, it becomes a sum of 3^m terms, each term involving m (iterated) single or double integrations. Three effectively different types of terms occur. Reading from right to left, single integrations with respect to the same variable throughout may occur (first type); or at least one single integration with respect to each of the two variables may occur before a double integration (second type); or a double integration will occur before single integrations with respect to each of the two variables (third type). In absolute value

$$\text{a term of the first type} \leqq K^m(\varphi_x + \varphi_y) \frac{(x - a + y - c)^{m-1}}{(m - 1)!}$$

$$\text{a term of the second type} \leqq K^m \cdot \varphi \frac{(x - a + y - c)^{m-2}}{(m - 2)!}$$

a term of the third type

$$\leqq \max.\left\{K^m \varphi \cdot L^{\frac{1}{2}} \frac{(x - a + y - c)^{m-2}}{(m - 2)!}, K^m \varphi \frac{(x - a + y - c)^{m-1}}{(m - 1)!}\right\},$$

where L is greater than the length of any interval lying in S. Thus every term is less, in absolute value, than

$$K^m(\varphi_x + \varphi_y + \varphi) \frac{(x - a + y - c)^{m-2}}{(m - 2)!} L'$$

where $L' = \max. (L, 1)$, and

$$| f_m(x, y) | \leqq 3^m \cdot K^m \cdot L' \cdot (\varphi_x + \varphi_y + \varphi) \cdot \frac{(x - a + y - c)^{m-2}}{(m - 2)!},$$

$$\sum_{m=0}^{\infty} | f_m(x, y) | \leqq M \cdot (\varphi_x + \varphi_y + \varphi),$$

where M is a constant. Thus $f(x, y)$ is the sum of a series which converges (absolutely) almost everywhere and $| f(x, y) | \leqq M(\varphi_x + \varphi_y + \varphi)$ implying that f is L.s.s.

Since B, C are bounded and E is L.s.s.,

$$\int_a^x C(w, y)f(w, y)\,dw \sim \sum_{m=0}^{\infty} \int_a^x C(w, y)f_m(w, y)\,dw,$$

$$\int_c^y B(x, z)f(x, z)\,dz \sim \sum_{m=0}^{\infty} \int_c^y B(x, z)f_m(x, z)\,dz,$$

$$\int_a^x \int_c^y E(w, z)f(w, z)\,dz\,dw \sim \sum_{m=0}^{\infty} \int_a^x \int_c^y E(w, z)f_m(w, z)\,dz\,dw.$$

It follows that f is a L.s.s. solution of the given equivalence.

If f^1, f^2 are both L.s.s. solutions of the given equivalence, the difference $f^3 = f^1 - f^2$ will satisfy

$$f^3(x, y) \sim \int_a^x C(w, y)f^3(w, y)\,dw + \int_c^y B(x, z)f^3(x, z)\,dz + \int_a^x \int_c^y E(w, z)f^3(w, z)\,dz\,dw.$$

Substituting for f^3 from this equivalence into the right side of the same equivalence and iterating this process, we can express f^3 as a sum of 3^m terms each involving m iterated integrations. We set

$$f_x = \int_c^{\delta(x)} |f^3(x, z)|\,dz \qquad f_y = \int_a^{\beta(y)} |f^3(w, y)|\,dw$$

$$f = \max.\left\{\int\int_S |f^3(w, z)|\,dz\,dw, \left(\int\int_S |f^3(w, z)|^2\,dz\,dw\right)^{\frac{1}{2}}\right\}$$

and obtain, by familiar evaluations,

$$|f^3(x, y)| \leq \frac{3^m K^m (L')^m (f_x + f_y + f)}{(m-2)!}$$

for $m = 3, 4, \cdots$. Since $f_x + f_y + f$ is finite almost everywhere, $f^3 \sim 0$, that is, $f' \sim f^2$, and there is exactly one L.s.s. solution of the equivalence given in the statement of the lemma.

PROOF OF THEOREM 3.2. Suppose that $(T^{(3)})^*g = g^*$. From condition (1) page 905, and Lemma 3.2, the functions g_x, g_y must be defined almost everywhere and be summable over I and

$$\bar{A}g_x + \bar{A}_x g + \int_y^1 (\bar{B}g_x + \bar{B}_x g) - \bar{C}g - \int_y^1 \bar{E}g \sim -\int_y^1 g^*.$$

Due to the restrictions imposed on A, B, C, E, the last equivalence implies that g_{xy} is defined almost everywhere and

$$\text{(III)} \quad \bar{A}g_{xy} + \bar{A}_y g_x + \bar{A}_x g_y + \bar{A}_{xy} g - \bar{B}g_x - \bar{B}_x g - \bar{C}g_y - \bar{C}_y g + \bar{E}g \sim g^*.$$

Replace $g(x, y)$ by an equivalent function which is abs. cont. in x for every y, and let $j(x, y)$ be a function which is abs. cont. in y and equivalent to g_x. From the preceding equivalence (III), j_y must be summable over I and

$$g(x, y) = g(0, y) + \int_0^x g_x(w, y)\,dw \sim g(0, y) + \int_0^x \left\{ j(w, 0) + \int_0^y j_y(w, z)\,dz \right\} dw$$

shows that $j(x, 0)$ is summable linearly in x. Since $g(x, y)$ (and with it, $g(0, y)$) is equivalent to a function abs. cont. in y, we can replace g by an equivalent function of the form

$$g(x, y) = g(0, y) + g(x, 0) - g(0, 0) + \int_0^x \int_0^y g_{xy}(w, z)\,dz\,dw$$

where $g(0, y)$, $g(x, 0)$ are abs. cont. in y and in x respectively, and g_{xy} is summable over I.

With this $g(x, y)$ the equivalences (1), (2), (3), page 905, become equalities. From (1) we obtain

$$\overline{A(x, 1)}\,g(x, 1) + \int_x^1 \overline{C(w, 1)}\,g(w, 1)\,dw = 0,$$

that is

$$f(x) = \int_x^1 p(w) f(w)\,dw$$

where

$$f(x) \equiv \overline{A(x, 1)}\,g(x, 1) \qquad |p(x)| \equiv \left| \frac{C(x, 1)}{A(x, 1)} \right| \leqq K < \infty.$$

Applying the usual methods of evaluation,

$$|f(x)| \leqq \left(\max_{0 \leqq w \leqq 1} |f(w)| \right) \frac{K^{m+1}}{(m+1)!}$$

for $m = 1, 2, \cdots$. Hence $f(x) \equiv 0$. Since $|A(x, 1)| > 0$, this implies $g(x, 1) \equiv 0$. Similarly $g(1, y) \equiv 0$. Using (2), (3), (4), we obtain in the same way $g(x, 0) \equiv 0$ and $g(0, y) \equiv 0$. Thus $g(x, y) = \int_0^x \int_0^y g_{xy}(w, z)\,dz\,dw$ and vanishes on the boundary of I.

Applying Lemma 3.3 to the equivalence (III), we see that g_{xy} is L.s.s., g is in $D^{(3)}$, and $T^{(3)}g = (T^{(3)})^*g$. This proves that $(T^{(3)})^* \subset T_0^{(3')}$. But conversely, if $T_0^{(3')}g$ is defined, then integration by parts shows that $\int_0^1 \int_0^1 T^{(3)}f \cdot \bar{g} = \int_0^1 \int_0^1 f \cdot \overline{T_0^{(3')}g}$ for all f in $D^{(3)}$, which means that $T_0^{(3')} \subset (T^{(3)})^*$. Hence $(T^{(3)})^* = T_0^{(3')}$, proving the part of the theorem concerning $(T^{(3)})^*$.

f is in $D_0^{(3)}$ if and only if $f(x, y) = \int_0^x \int_0^y \varphi(w, z)\,dz\,dw$, where φ is L.s.s. and

orthogonal to every L.s.s. function ψ which is a function of x or y alone. The condition

$$\int_0^1\int_0^1 T^{(3)}f\cdot\bar{g} = \int_0^1\int_0^1 f\cdot\overline{g^*}$$

for all f in $D^{(3)}$, leads to the equivalent condition (use the above representation for f),

$$\bar{A}g + \int_y^1 \bar{B}g + \int_x^1 \bar{C}g + \int_x^1\int_y^1 \bar{E}g \sim \int_x^1\int_y^1 g^* + h,$$

where h may be any function which is in the smallest closed linear subspace of H containing all L.s.s. functions $\psi(x)$, $\psi(y)$ (such an $h(x, y)$ must be of the form $h_1(x) + h_2(y)$ with h_1, h_2 L.s.s., and any function of this form will be possible). We now apply Lemma 3.4 to show that

$$\bar{A}g_1 + \int_y^1 \bar{B}g_1 + \int_x^1 \bar{C}g_1 + \int_x^1\int_y^1 \bar{E}g_1 \sim h$$

has a unique L.s.s. solution g_1. Then we have

$$\bar{A}(g - g_1) + \int_y^1 \bar{B}(g - g_1) + \int_x^1 \bar{C}(g - g_1) + \int_x^1\int_y^1 \bar{E}(g - g_1) \sim \int_x^1\int_y^1 g^*,$$

and an argument similar to that used in determining $(T^{(3)})^*$ shows that

$$g_2(x, y) \equiv g(x, y) - g_1(x, y) = \int_x^1\int_y^1 \varphi(w, z)\,dz\,dw$$

for some L.s.s. φ, and

$$g^* = (\bar{A}g_2)_{xy} - (\bar{B}g_2)_x - (\bar{C}g_2)_y + (\bar{E}g_2).$$

This completes the proof of Theorem 3.2.

Let $f = f_1 + f_2$, $g = g_1 + g_2$ and

(i) $$\overline{A^*(x, y)}f_1(x, y) - \int_0^y \overline{B^*(x, z)}f_1(x, z)\,dz - \int_0^y \overline{C^*(w, y)}f_1(w, y)\,dw + \int_0^x\int^y \overline{E^*}(w, z)f(w, z)\,dz\,dw \sim k_1(x) + k_2(y),$$

(ii) $$f_2(x, y) = \int_0^x\int_0^y (f_2)_{xy}(w, z)\,dz\,dw,$$

(iii) $$\overline{A(x, y)}\,g_1(x, y) + \int_y^1 \overline{B(x, z)}\,g_1(x, z)\,dz + \int_x^1 C(w, y)\,g_1(w, y)\,dw + \int_x^1\int_y^1 E(w, z)\,g_1(w, z)\,dz\,dw \sim h_1(x) + h_2(y),$$

(iv) $$g_2(x, y) = \int_x^1\int_y^1 (g_2)_{xy}(w, z)\,dz\,dw,$$

with k_1, k_2, $(f_2)_{xy}$, h_1, h_2, $(g_2)_{xy}$, L.s.s. on I. Integration by parts gives

$$\int\int_I \{T^{(3)}f\cdot\bar{g} - f\cdot\overline{T^{(3')}g}\} = \int\int_I [(f_2)_{xy}(\overline{h_1 + h_2}) - (\bar{g}_2)_{xy}(k_1 + k_2)].$$

By means of this identity we could discuss linear boundary conditions.[36]

Theorems 3.1, 3.2 can be extended to the general region S, and by a transformation of variables, to a more general form of operational expression. Let the differential equation

$$A(x, y)\,dy^2 + 2B(x, y)\,dx\,dy + C(x, y)\,dx^2 = 0$$

define two real systems of curves Γ_1, Γ_2, in a finite region R, such that

(i) through each point of R passes exactly one curve of each system.

(ii) there are two fixed curves γ_1, γ_2, one in each of the systems Γ_1, Γ_2, which meet all curves of the other system in exactly one point of R.

Then Theorems 3.1, 3.2 can be generalized to an operational expression

$$A\frac{\partial^2}{\partial x^2} + 2B\frac{\partial^2}{\partial x\,\partial y} + C\frac{\partial^2}{\partial y^2} + 2D\frac{\partial}{\partial x} + 2E\frac{\partial}{\partial y} + F$$

on the region R, the statement of Theorem 3.2 involving linear integration along the characteristic curves Γ_1, Γ_2.

4. The Elliptic Differential Operator

The following lemma will be required in this section.

Lemma 4.1. *Let $f_1(x, y)$, $f_2(x, y)$ be defined and equivalent on a finite rectangle, and L.s.s. If*

(i) *f_1, $(f_1)_x$ are abs. cont. in x for fixed y, $0 < y < 2\pi$, and, $(f_1)_{xx}$, and necessarily with it, $(f_1)_x$, are L.s.s.*

(ii) *f_2, $(f_2)_y$ are abs. cont. in y for fixed x, $0 < x < 2\pi$, and $(f_2)_{yy}$, and necessarily with it, $(f_2)_y$, are L.s.s.*

then there is a function $f(x, y)$ equivalent to f_1 and to f_2, such that $f(x, y)$ is continuous on the open rectangle and satisfies (i) *and* (ii).[37]

Proof. We may assume, by a linear transformation of variables, that the rectangle is I_π: $0 \leqq x \leqq 2\pi$; $0 \leqq y \leqq 2\pi$. Define $\theta(x, y)$ so that

(i) θ, θ_x, θ_y, are continuous on the closed square I_π.

(ii) θ, θ_x, are abs. cont. in x and vanish for $x = 0$, $x = 2\pi$.

---

[36] If only functions of a sufficiently high order of differentiability are involved, then the right hand side of the identity can be partially integrated into the classical form of a boundary condition. That this process can really fail in the general case is shown by the simple special case of the expression $\partial^2/\partial x\partial y$: the closure of $T^{(3)}$ will be defined here for every L.s.s. function of the form $h(x) + k(y)$, and the h, k need not be differentiable in the usual sense nor even continuous.

[37] In this section, the symbol f_x will denote the partial derivative of the function f itself, not merely of a suitable equivalent function.

(iii) θ, θ_y, are abs. cont. in y and vanish for $y = 0$, $y = 2\pi$.
(iv) θ_{xx}, θ_{yy} are effectively bounded on I_π.
(v) $\theta(x, y) = 1$ for $\epsilon \leqq x, y \leqq 2\pi - \epsilon$ $(0 < \epsilon < 1)$.[38]
Then

$f_1\theta$ is L.s.s. over I_π, and vanishes for $x = 0$ and for $x = 2\pi$.
$(f_1\theta)_x \sim f_1\theta_x + (f_1)_x\theta$ is L.s.s. over I_π and vanishes for $x = 0$ and for $x = 2\pi$.
$(f_1\theta)_{xx} \sim f_1\theta_{xx} + 2(f_1)_x\theta_x + (f_1)_{xx}\theta$ is L.s.s. over I_π.

If $f_1\theta \sim f_2\theta \sim \sum_{m,n} a_{mn}e^{imx+iny}$ it will follow that $\sum_{m,n} m^4 \mid a_{mn} \mid^2$ is finite. Similarly $\sum_{m,n} n^4 \mid a_{mn} \mid^2$ is finite and

$$\sum_{m^2+n^2\neq 0} \mid a_{mn} \mid \leqq \left(\sum_{m^2+n^2\neq 0} (m^2 + n^2)^2 \mid a_{mn} \mid^2\right)^{\frac{1}{2}} \left(\sum_{m^2+n^2\neq 0} \frac{1}{(m^2 + n^2)^2}\right)^{\frac{1}{2}}$$

is also finite, showing that $\sum_{m,n} a_{mn}e^{imx+iny}$ is absolutely convergent and hence continuous. Thus $f_1(x, y)$ is equivalent to a continuous function on $\epsilon \leqq x, y \leqq 2\pi - \epsilon$, for arbitrarily small ϵ, and hence on the open rectangle I_π. If we set $f(x, y)$ equal to this continuous function for $0 < x, y < 2\pi$ and

$$f(x, 0) = f_2(x, 0) \qquad f(0, y) = f_1(0, y)$$
$$f(x, 2\pi) = f_2(x, 2\pi) \qquad f(2\pi, y) = f_1(2\pi, y)$$

then this $f(x, y)$ will satisfy the requirements of the lemma.

In this §, H will be the space of functions L.s.s. on a finite rectangle which we may take as I_π without loss of generality. The differential expression

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

can be applied directly to the functions $f(x, y)$ in the linear domains, dense in H, defined by the

Definition 4.1. $D^{(4)}$ consists of the f in H with the following properties:
(i) f_1, $(f_1)_x$ are abs. cont. in x, and $(f_1)_{xx}$ is in H, for some $f_1 \sim f$.
(ii) f_2, $(f_2)_y$ are abs. cont. in y, and $(f_2)_{yy}$ is in H, for some $f_2 \sim f$.

---

[38] For instance, define

$$\theta(x, y) \begin{cases} = 1 & \text{for } \epsilon \leqq x,\ y \leqq 2\pi - \epsilon \\ = \dfrac{x^2(2\epsilon - x)^2}{\epsilon^4} & \text{for } \begin{cases} 0 \leqq x \leqq \epsilon \text{ and} \\ \epsilon \leqq y \leqq 2\pi - \epsilon \end{cases} \\ = \dfrac{2\epsilon(x + y) - x^2 - y^2 - \epsilon^2}{\epsilon^4} & \text{for } \begin{cases} 0 \leqq x, y \leqq \epsilon \text{ and} \\ (\epsilon - x)^2 + (\epsilon - y)^2 \leqq \epsilon^2. \end{cases} \\ = 0 & \text{for } \begin{cases} 0 \leqq x, y \leqq \epsilon \text{ and} \\ (\epsilon - x)^2 + (\epsilon - y)^2 \geqq \epsilon^2 \end{cases} \end{cases}$$

and set $\theta(x, y) = \theta(y, x) = \theta(2\pi - x, y) = \theta(2\pi - x, 2\pi - y)$.

REMARK. By Lemma 4.1 we may assume that $f \equiv f_1 \equiv f_2$.

DEFINITION 4.2. $D_\tau^{(4)}$ consists of the f in $D^{(4)}$ for which

$$f(0, y) = f(2\pi, y) \qquad f_x(0, y) = f_x(2\pi, y)$$
$$f(x, 0) = f(x, 2\pi) \qquad f_y(x, 0) = f_y(x, 2\pi).$$

DEFINITION 4.2. $D_0^{(4)}$ consists of the f in $D_\tau^{(4)}$ for which $f(0, y) = f(x, 0) = f_x(0, y) = f_y(x, 0) = 0$.

The operators determined by $\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$ on the above defined domains, will be denoted by $T^{(4)}$, $T_\tau^{(4)}$, and $T_0^{(4)}$ resp.

THEOREM 4.1. *$T_\tau^{(4)}$ is a self-adjoint operator, that is $(T_\tau^{(4)})^* = T_\tau^{(4)}$.*

PROOF. If $\displaystyle\int\int_{I_\pi} \{(f_{xx} + f_{yy})\bar{g} - f\overline{g^*}\} = 0$ for all f in $D_\tau^{(4)}$, in particular for $f = e^{imx+iny}$ $(m, n = 0, \pm 1, \pm 2, \cdots)$ it follows that if $g \sim \sum_{m,n} g_{mn} e^{imx+iny}$, then $g^* \sim \sum_{m,n}(-m^2 - n^2) g_{mn} e^{imx+iny}$ and $\sum_{m,n}(m^2 + n^2)^2 |g_{mn}|^2$ is finite. By the inequality of Schwarz,

$$\sum_{m^2+n^2\neq 0} |g_{mn}| \leqq \left\{\sum_{m^2+n^2\neq 0} (m^2 + n^2)^2 |g_{mn}|^2\right\}^{\frac{1}{2}} \left\{\sum_{m^2+n^2\neq 0} \left(\frac{1}{m^2 + n^2}\right)^2\right\}^{\frac{1}{2}}$$

is finite. Thus $g(x, y)$ is the sum of an absolutely convergent exponential series, and hence is in $D_\tau^{(4)}$, and $(T_\tau^{(4)})^* g = T_\tau^{(4)} g$. This means that $(T_\tau^{(4)})^* \subset T_\tau^{(4)}$.

Conversely, if $T_\tau^{(4)} g$ is defined, then

$$\int\int_{I_\pi} T_\tau^{(4)} f \cdot \bar{g} = \int\int_{I_\pi} (f_{xx} \cdot \bar{g} + f_{yy} \cdot \bar{g}) = \int\int_{I_\pi} (f \cdot \bar{g}_{xx} + f \cdot \bar{g}_{yy}) = \int\int_{I_\pi} f \cdot \overline{T_\tau^{(4)} g},$$

that is, $(T_\tau^{(4)})^* g$ is defined and equals $T_\tau^{(4)} g$. This means that $T_\tau^{(4)} \subset (T_\tau^{(4)})^*$, and hence $(T_\tau^{(4)})^* = T_\tau^{(4)}$.

THEOREM 4.2. *$(T^{(4)})^* = T_0^{(4)}$ and hence $(T_0^{(4)})^* = \bar{T}^{(4)}$.*

PROOF. Since $T^{(4)} \supset T_\tau^{(4)}$ we have $(T^{(4)})^* \subset (T_\tau^{(4)})^*$. Thus $(T^{(4)})^* g = g^*$ will imply that g is in $D_\tau^{(4)}$, $g^* = g_{xx} + g_{yy}$ and

$$\int\int_{I_\pi} (f_{xx} + f_{yy})\bar{g} = \int\int_{I_\pi} f \cdot \overline{(g_{xx} + g_{yy})}$$

for every f in $D^{(4)}$. Setting $f(x, y) = y \cdot \displaystyle\int_0^x (x - z) f''(z)\, dz$, where $f''(z)$ is L.s.s. linearly over $0 \leqq z \leqq 2\pi$, we obtain

$$\int\int_{I_\pi} y f''(x)\overline{g(x, y)}\, dy\, dx = \int\int_{I_\pi} y \cdot \left(\int_0^x (x - z) f''(z)\, dz\right) \overline{(g_{xx}(x, y) + g_{yy}(x, y))}\, dy\, dx,$$

$$\int_0^{2\pi} f''(x)\, dx \int_0^{2\pi} y \cdot \overline{g(x, y)}\, dy = \int_0^{2\pi} \left(\int_0^x (x - z) f''(z)\, dz\right) dx \int_0^{2\pi} y \overline{g_{yy}(x, y)}\, dy$$
$$+ \int_0^{2\pi} y\, dy \int_0^{2\pi} \left(\int_0^x (x - z) f''(z)\, dz\right) \overline{g_{xx}(x, y)}\, dx,$$

$$\int_0^{2\pi} f''(x) \left\{ \int_0^{2\pi} y \cdot \overline{g(x, y)}\, dy - 2\pi \int_0^{2\pi} (z - x)\overline{g_y(z, 2\pi)}\, dz \right.$$

$$\left. - \int_0^{2\pi} y \cdot dy \int_x^{2\pi} (z - x)\overline{g_{xx}(z, y)}\, dz \right\} dx = 0$$

for all L.s.s. f''. Hence the expression in the bracket of the last equality is equivalent to zero. That is,

$$2\pi \int_x^{2\pi} (z - x) g_y(z, 2\pi)\, dz + (2\pi - x) \int_0^{2\pi} y \cdot g_x(2\pi, y)\, dy - \int_0^{2\pi} y \cdot g(2\pi, y)\, dy \sim 0$$

($\sim$ in x). Both sides of the previous equivalence are abs. cont. in x and differentiation gives

$$2\pi \int_x^{2\pi} g_y(z, 2\pi)\, dz = -\int_0^{2\pi} y \cdot g_x(2\pi, y)\, dy = 2\pi \int_{2\pi}^{2\pi} g_y(z, 2\pi)\, dz = 0.$$

Thus $g_y(x, 2\pi) = 0$. Similarly $g_y(x, 0) = g_x(2\pi, y) = g_x(0, y) = 0$.

In the same way, starting with $f(x, y) = y^2 \int_0^x (x - z) f''(z)\, dz$, we obtain $g(x, 0) = g(x, 2\pi) = g(0, y) = g(2\pi, y) = 0$.

This implies that g is in $D_0^{(4)}$ and $(T^{(4)})^* g = T_0^{(4)} g$.

Conversely, if $T_0^{(4)} g$ is defined, then $\int\int_{I_\pi} (T^{(4)} f \cdot \bar{g} - f \cdot \overline{T_0^{(4)} g}) = 0$ for every f in $D^{(4)}$, as can be verified by integration by parts. Consequently $(T^{(4)})^* = T_0^{(4)}$.

THEOREM 4.3. *$\tilde{T}^{(4)} = (T_0^{(4)})^*$ and is defined for exactly those f which have the form $f_1 + f_2$ where*

(i) f_1 is harmonic and L.s.s. on I_π and is the limit of a sequence of harmonic functions from $D^{(4)}$.

(ii) f_2 is in $D^{(4)}$.

The value of the closure is given by $(\tilde{T}^{(4)}) f = T^{(4)} f_2$.

PROOF. The sufficiency of these conditions is clear. The necessity can be shown as follows: If $(\tilde{T}^{(4)}) f = g$, there is a sequence f_m in $D^{(4)}$ such that

$$f - f_m, \qquad g - T^{(4)} f_m,$$

converge to zero simultaneously as m tends to infinity. We can put f_m in the form

$$f_m(x, y) = f_{m,1}(x, y) + \frac{1}{4\pi} \int\int_{I_\pi} \left\{ \log\left[(x - \xi)^2 + (y - \eta)^2\right] \right\} \left\{ \frac{\partial^2 f_m}{\partial \xi^2}(\xi, \eta) \right.$$

$$\left. + \frac{\partial^2}{\partial \eta^2} f_m(\xi, \eta) \right\} d\xi\, d\eta$$

where $f_{m,1}(x, y)$ is a harmonic function, L.s.s. on I_r. By a theorem of Lichtenstein,[39] $(f_m - f_{m,1})$, and with it, $f_{m,1}$, are in $D^{(4)}$. By the inequality of Schwarz

$$\left|\frac{1}{4\pi}\int\int_{I_r}\left\{\log\left[(x-\xi)^2+(y-\eta)^2\right]\right\}\left\{\frac{\partial^2}{\partial\xi^2}f_m(\xi,\eta)+\frac{\partial^2}{\partial\eta^2}f_m(\xi,\eta)-g(\xi,\eta)\right\}d\xi\,d\eta\right|^2$$

$$\leqq \frac{1}{16\pi^2}\int\int_{I_r}\left|\log\left[(x-\xi)^2+(y-\eta)^2\right]\right|^2 d\xi\,d\eta\cdot\int\int_{I_r}\left|\frac{\partial^2}{\partial\xi^2}f_m(\xi,\eta)\right.$$

$$\left.+\frac{\partial^2 f_m}{\partial\eta^2}(\xi,\eta)-g(\xi,\eta)\right|^2 d\xi\,d\eta$$

converges uniformly in x, y to zero as m tends to infinity. It follows that $f_{m,1}(x, y)$ converges in H to a limit $f_1(x, y)$. Since the limit (in H) of a sequence of harmonic functions is necessarily harmonic, $f(x, y)$ has the form

$$f(x,y)=f_1(x,y)+\frac{1}{4\pi}\int\int_{I_r}\left\{\log\left[(x-\xi)^2+(y-\eta)^2\right]\right\}\cdot g(\xi,\eta)\,d\xi\,d\eta$$

and

$$g(\xi,\eta)\sim\frac{\partial^2 f}{\partial\xi^2}(\xi,\eta)+\frac{\partial^2 f}{\partial\eta^2}(\xi,\eta).$$ [39]

Remark. Theorem 4.3 shows that for the elliptic operator, just as for the hyperbolic operator, the closure is effectively obtained by specifying precisely for which functions the closure operator shall vanish. For the hyperbolic operator we have succeeded in doing this by means of line integrals along the characteristic curves, and we found that the differential expression can not always be applied directly to these functions. For the elliptic operator, however, we know that the differential expression will apply directly to these functions (they are harmonic), but we have not succeeded in characterizing them explicitly.

Princeton University.

---

[39] L. Lichtenstein, Crelle's Journal 141 (1912) pp. 34–42, proves that if $h(x, y)$ is L.s.s., then

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\int\int_{I_r}\left\{\log\left[(x-\xi)^2+(y-\eta)^2\right]\right\}h(\xi,\eta)\,d\xi\,d\eta$$

exists almost everywhere and is equivalent to $4\pi h(x, y)$.

Annals of Mathematics
Vol. 38, No. 4, October, 1937

AN EXAMPLE OF A NON-CLOSED CONNECTED SUBGROUP OF THE TWO-DIMENSIONAL VECTOR SPACE

By E. Livenson

(Received October 6, 1936)

In his very interesting work entitled: "Topologische Gruppen mit genügend vielen fastperiodischen Funktionen" (Ann. of Math. 37 (1936), p. 157–77), among many other assertions Freudenthal states (without proof) the following theorem:

Every connected subgroup of a vector space is closed (see p. 59).

1. Nevertheless I shall construct (using Zermelo's Axiom) in two-dimensional vector space an example of a subgroup which is connected but not closed. To this end I shall construct first of all a Hamel's base[1] of a special type.

2. We shall say that a set $\mathfrak{S}$ of real numbers consists of *linearly independent* numbers if between them exists no linear commensurability relation

$$\alpha_1 s_1 + \cdots + \alpha_n s_n = 0,$$

where $s_i \epsilon \mathfrak{S}$ $(i = 1, \cdots, n)$ and α_i are rational numbers not all zero. Now let

$$P_0, P_1, P_2, \cdots, P_\xi, \cdots \mid \Omega_\gamma,$$

where Ω_γ is the least ordinal of the cardinal number of the continuum, be a transfinite sequence consisting of all plane perfect sets.[2] Choose two linearly independent real numbers u' and u''. We shall define presently for every $\xi < \Omega_\gamma$ two numbers $V_\xi^{(1)}$ and $V_\xi^{(2)}$ in such a way that the set $\mathfrak{S}_{\Omega_\gamma}$ of all numbers $u', u'', V_\xi^{(1)}, V_\xi^{(2)}$ $(\xi = 1, 2, \cdots \mid \Omega_\gamma)$ constitutes a Hamel's base.

3. We shall define these numbers inductively. Suppose that for every $\xi < \xi_0$ the numbers $V_\xi^{(1)}$ and $V_\xi^{(2)}$ are already defined, and that every set $\mathfrak{T}_\xi$ of all numbers $u', u'', V_\eta^{(1)}, V_\eta^{(2)}$ $(\eta \leqq \xi)$ consists of linearly independent numbers. Then evidently the set $\mathfrak{S}_{\xi_0} = \sum_{\xi < \xi_0} \mathfrak{T}_\xi$ consists of linearly independent numbers. Denote by $\bar{\mathfrak{S}}_{\xi_0}$ the set of all the points (x, y) of the plane such that both x and y are linearly dependent on $\mathfrak{S}_{\xi_0}$. The cardinal number of $\bar{\mathfrak{S}}_{\xi_0}$ being evidently less than $\mathfrak{c}$ there exists a point

$$(x_0, y_0) \epsilon P_{\xi_0} - \bar{\mathfrak{S}}_{\xi_0}. \tag{1}$$

One of two possibilities presents itself:

(a) Either there exists no commensurability relation

$$\alpha x_0 + \beta y_0 = \mu u' + \nu u'' + S, \tag{2}$$

[1] G. Hamel, Math. Annalen, 60, (1905), pp. 459–462.

[2] Compare with Hausdorff, *Mengenlehre*, 3rd. ed., p. 176.

where α, β, μ, ν are rational numbers not all zero and S is a linear form in $V_\xi^{(1)}$, $V_\xi^{(2)}$ ($\xi < \xi_0$) with rational coefficients; in this case let

$$V_{\xi_0}^{(1)} = x_0 \, , \; V_{\xi_0}^{(2)} = y_0. \tag{3}$$

Or

(b) There exists a commensurability relation (2). We can find then three rational numbers λ_1, λ_2, λ_3 satisfying two equations

$$\begin{aligned} \alpha\lambda_1 + \beta\lambda_2 &= \mu, \\ \alpha\lambda_2 + \beta\lambda_3 &= \nu. \end{aligned} \tag{4}$$

One at least of the numbers α, β is not equal to 0. Suppose, e.g., that it is β. Then x_0 is linearly independent of $\mathfrak{S}_{\xi_0}$ (because otherwise x_0 and y_0 would be both linearly dependent on $\mathfrak{S}_{\xi_0}$ in contradiction with (1)). Let

$$V_{\xi_0}^{(1)} = x_0 - \lambda_1 u' - \lambda_2 u''. \tag{5}$$

$V_{\xi_0}^{(1)}$ is linearly independent of $\mathfrak{S}_{\xi_0}$.
Combining (2), (4) and (5) we obtain

$$y_0 = \lambda_2 u' + \lambda_3 u'' - \frac{\alpha}{\beta} V_{\xi_0}^{(1)} + \frac{1}{\beta} S. \tag{6}$$

If $\beta = 0$, $\alpha \neq 0$, we define in the same manner

$$V_\xi^{(1)} = y_0 - \lambda_2 u' - \lambda_3 u''$$

and then

$$x_0 = \lambda_1 u' + \lambda_2 u'' + \frac{1}{\alpha} S - \frac{\beta}{\alpha} V_{\xi_0}^{(1)}.$$

In both cases let $V_{\xi_0}^{(2)}$ be an arbitrary number linearly independent of the set $\mathfrak{S}_{\xi_0} + (V_{\xi_0}^{(1)})$. In case a) as well as b) the set $\mathfrak{T}_{\xi_0} = \mathfrak{S}_{\xi_0} + (V_{\xi_0}^{(1)}) + (V_{\xi_0}^{(2)})$ consists of linearly independent numbers.

4. It is evident that the set $\mathfrak{S}_{\Omega_\gamma}$ consists of linearly independent numbers. Besides, comparing (3), (5), (6) we see that it possesses the following property: Every plane perfect set P contains a point (x_0, y_0) such that

$$x_0 = \lambda_1 u' + \lambda_2 u'' + \alpha_1 V_{\xi_1}^{(i_1)} + \cdots + \alpha_k V_{\xi_k}^{(i_k)}$$

($i_s = 1, 2$; $\xi_s < \Omega_\gamma$),

$$y_0 = \lambda_2 u' + \lambda_3 u'' + \beta_1 V_{\eta_1}^{(j_1)} + \cdots + \beta_l V_{\eta_l}^{(j_l)}$$

($j_s = 1, 2$; $\eta_s < \Omega_\gamma$), where λ_1, λ_2, λ_3, α_s, β_s are rational numbers. In particular if P is a segment of the line $x = x_0$, we see that every real number x_0 can be represented as a linear form in the elements of $\mathfrak{S}_{\Omega_\gamma}$ with rational coefficients. This together with the linear independence of the elements of $\mathfrak{S}_{\Omega_\gamma}$ proves that $\mathfrak{S}_{\Omega_\gamma}$ is a Hamel's base.

5. $\mathfrak{S}_{\Omega_\gamma}$ being a Hamel's base, every real number can be expressed, and in one

way only, as a linear form in its elements with rational coefficients. Consider the set R of all points (x, y) such that the coefficient of u'' in x is equal to the coefficient of u' in y. It is obvious that this set is a subgroup of the vector space $\{(x, y)\}$ and that it is not closed. We shall prove now that it is connected.

6. Suppose the contrary and let $R_1 \subset R (0 \neq R_1 \neq R)$ be closed and open in R simultaneously. R_1 being open in R there exists an open plane set V such that $R_1 = RV$. R being everywhere dense in the plane we have

$$R \cdot \bar{R}_1 = R \cdot \overline{RV} = R \cdot \bar{V}$$

But R_1 is closed in R i.e. $R \cdot \bar{R}_1 = R \cdot \bar{V} = R_1$ or

$$R \cdot (\bar{V} - V) = 0. \tag{7}$$

On the other hand V cannot contain all but a countable number of points (because $R - R_1$ is open in R and non vacuous). Therefore the set $\bar{V} - V$ contains a perfect subset. But we have seen (in Art. 4) that every perfect set contains a point of R, which contradicts (7). We have arrived at a contradiction, which proves that our original supposition (that R is not connected) is false. Consequently R is connected, q.e.d.

LENINGRAD STATE UNIVERSITY, U. S. S. R.

ANNALS OF MATHEMATICS
Vol. 38, No. 4, October, 1937

NORMALIZED INTEGRAL BASES OF ALGEBRAIC NUMBER FIELDS I[1]

BY A. ADRIAN ALBERT

(Received January 8, 1937)

INTRODUCTION

Every algebraic number field $\mathfrak{F}$ of degree n over the rational number field is defined by a root θ of an irreducible rational equation $f(x) = 0$ of degree n. One of the most interesting problems in the theory of such fields is that of finding explicit formulae for the quantities of a basis of the integers of $\mathfrak{F}$ in terms of θ and the coefficients of $f(x)$. This problem has had no general solution and has not been completely solved even for the case[2] $n = 4$.

Some simplifications in the above formulae may be obtained by choosing θ to be an algebraic integer and such that $a\theta$ is an algebraic integer for rational a if and only if a is integral.[3] These are partial normalizations on θ. Additional simplifications of the formulae are obtained when we choose θ to have zero trace,[4] that is the coefficient of x^{n-1} in $f(x)$ to be zero.

In the present paper we shall obtain simplifications by making further normalizations on θ. We shall in fact show that every $\mathfrak{F}$ of degree $n > 2$ is generated by a quantity θ such that in addition to the above properties the first two elements of an integral basis of $\mathfrak{F}$ are

$$1, \theta.$$

This result will be completed here by our giving necessary and sufficient conditions on $f(x)$ that θ have the desired properties. The special cases $n = 3, 4, 5$ are of particular interest and we shall apply our formal conditions to obtain more explicit results in these cases.

In the theory of cubic fields our above normalization shows that

$$1, \theta, \Omega = \frac{d + c\theta + \theta^2}{E}$$

are an integral basis of $\mathfrak{F}$, where Ω is to be determined. Certain known results[4] reduce the problem of finding Ω to the solution of the congruences

$$f(e) \equiv 0(E^2), \qquad f'(e) \equiv 0(E),$$

[1] Presented to the Society December 29, 1936.

[2] The second part of this paper will be devoted to new results in the theory of quartic fields.

[3] See Berwick's *Integral Bases*.

[4] This normalization is customary in the case of quadratic fields and was made by the author in the theory of cubic fields, these *Annals*, vol. 31 (1930), pp. 550–566.

where $f'(x)$ is the derivative of $f(x)$. These results have never been generalized and the author has long been interested in their meaning. We shall show here that the generalization consists in a consideration of integral domains in $\mathcal{F}$ with a basis

$$1, \theta, \theta^2, \cdots, \theta^{n-2}, \Omega = \frac{e_1 + \cdots + e\theta^{n-2} + \theta^{n-1}}{E},$$

with the above pair of congruences again the sole conditions to be satisfied.

The author's results in his determination of the integers of all cubic fields[4] had many sub-cases. This was due to the fact that only the usual normalizations were employed. We shall employ the further normalization given above and shall use these results to obtain a completely normalized generation giving as simple a basis $1, \theta, \Omega$ as possible. Our final result will provide an invariantive classification of cubic fields, that is a classification into sets such that no field in any one set is equivalent to a field in any other. Moreover we shall give an explicit determination of the discriminant of any cubic field in terms of a normalized generation.

Chapter I. Fundamentals and Notations

1. **The terminology.** In our determination of normalized integral bases of algebraic number fields it will be frequently necessary to distinguish between ordinary (or rational) integers and algebraic integers. We shall accomplish this by introduction of a uniform notation.

Algebraic numbers will be designated by Greek letters. They may be rational. But any number designated by a Roman letter will necessarily be rational. This usage will be of particular importance in the case of integers.

All of the congruences appearing in our treatment will be ordinary integral congruences. We use the short notation

$$a \equiv b(c)$$

in the usual fashion to mean that a, b, c are rational integers, c divides $a - b$. If p is a rational integral prime such that p^a divides an integer b but p^{a+1} does not divide b we say that b is *exactly divisible* by p^a.

Sets of quantities will be designated by script capitals. The most frequent of these are the field $\mathcal{R}$ of all rational numbers, that is the coefficient field of our subject, the algebraic number field $\mathcal{F}$ whose integers we are studying, the integral domain $\mathcal{I}$ of all rational integers, and the integral domain $\mathcal{J}$ of all integers of $\mathcal{F}$.

2. **Algebraic number fields.** Let us recall some of the elementary concepts of the theory of algebraic numbers. They are well known to the reader but we shall give the results mainly to introduce our notations.

A complex number θ is called an *algebraic number* if it is a root of an equa-

tion with rational coefficients. Then θ is also a root of its unique *minimum equation*

(1) $$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n = 0 \qquad (a_i \text{ in } \mathfrak{R}),$$

which is irreducible in $\mathfrak{R}$. The field $\mathfrak{F} = \mathfrak{R}(\theta)$ of degree n over $\mathfrak{F}$ generated by θ is the set of all rational functions of θ with rational coefficients. Every number ϕ of $\mathfrak{F}$ is uniquely expressible in the form

(2) $$\phi = x_1 + x_2\theta + \cdots + x_n\theta^{n-1} \qquad (x_i \text{ in } \mathfrak{R}),$$

so that $1, \theta, \cdots, \theta^{n-1}$ are linearly independent in $\mathfrak{R}$. Any n quantities $\mu_1, \cdots, \mu_n$ of $\mathfrak{F}$ which are linearly independent in $\mathfrak{R}$ form a basis of $\mathfrak{F}$ over $\mathfrak{R}$, that is every ϕ of $\mathfrak{F}$ is uniquely expressible in the form

(3) $$\phi = y_1\mu_1 + \cdots + y_n\mu_n \qquad (y_i \text{ in } \mathfrak{R}).$$

The polynomial $f(x)$ has the factorization

(4) $$f(x) = (x - \theta^{(1)})(x - \theta^{(2)}) \cdots (x - \theta^{(n)})$$

with $\theta^{(1)} = \theta, \theta^{(2)}, \cdots, \theta^{(n)}$ ordinary complex numbers. Each $\theta^{(i)}$ generates an algebraic number field $\mathfrak{F}^{(i)} = \mathfrak{R}(\theta^{(i)})$ conjugate to $\mathfrak{F} = \mathfrak{R}(\theta)$ over $\mathfrak{F}$, that is equivalent to $\mathfrak{F}$ under the correspondence defined by (2) and

$$\phi \leftrightarrow \phi^{(i)} = x_1 + x_2\theta^{(i)} + \cdots + x_n(\theta^{(i)})^{n-1}.$$

Then the $\phi^{(i)}$ are called the *conjugates* of ϕ. It is easily seen that ϕ is a rational number (that is $x_2 = \cdots = x_n = 0$) if and only if $\phi = \phi^{(2)} = \cdots = \phi^{(n)}$.

The numbers ϕ of $\mathfrak{F}$ are themselves algebraic numbers. In fact ϕ is a root of

(5) $$g(y) = (y - \phi)(y - \phi^{(2)}) \cdots (y - \phi^{(n)}) = 0$$

with rational coefficients. The field $\mathfrak{R}(\phi)$ is a subfield of $\mathfrak{F}$ and has degree m over $\mathfrak{F}$ if and only if

(6) $$n = qm, \qquad g(y) = [h(y)]^q,$$

where the minimum function $h(y)$ of ϕ has degree m. The field $\mathfrak{R}(\phi)$ has degree n, $\mathfrak{R}(\phi) = \mathfrak{F}$, ϕ generates $\mathfrak{F}$ if and only if $q = 1$, $g(y)$ has no multiple roots. For, $h(y) = 0$ is the irreducible minimum equation of ϕ. We apply this to state

LEMMA 1. *A number ϕ of $R(\theta)$ generates $\mathfrak{F}$ if and only if the discriminant of ϕ*

(7) $$d(\phi) = \prod_{\substack{i=1,\cdots,n-1 \\ j=2,\cdots,n \\ i<j}} [\phi^{(i)} - \phi^{(j)}]^2$$

is not zero.

Our lemma states that $1, \phi, \phi^2, \cdots, \phi^{n-1}$ form a basis of $\mathfrak{F}$ over $\mathfrak{R}$ if and only if $d(\phi) \neq 0$. The particular case where

(8) $$\phi = u + v\theta \qquad (u, v \text{ in } \mathfrak{R}),$$

is a linear combination of 1 and θ evidently gives

$$d(\phi) = v^{2n} d(\theta). \tag{9}$$

But $f(x)$ is irreducible so that

$$d(\theta) \neq 0. \tag{10}$$

Hence ϕ of (8) generates $\mathfrak{F}$ if and only if $v \neq 0$.

3. **A lemma on integers.** In our study of normalized generations of $\mathfrak{F}$ over $R(\theta)$ we shall require certain elementary results on rational integers. Consider a set $g_1, \cdots, g_m$ of integers. Their greatest common divisor (abbreviated g.c.d.) is always expressible in the form[5]

$$g = g_1 h_1 + \cdots + g_m h_m > 0 \qquad (h_i \text{ in } \mathfrak{I}). \tag{11}$$

We call $g_1, \cdots, g_m$ a primitive set if $g = 1$. The case $m = 1$ states that g_1 is primitive if $g_1 = \pm 1$. Moreover $g_1, \cdots, g_m$ is always primitive if some $g_i = \pm 1$. We use these simple facts in a brief proof of the known but highly important.[6]

LEMMA 2. *Let $g_1, \cdots, g_m$ be a set of rational integers for $m > 1$. Then there exist integers g_{ij} such that*

$$\delta = \begin{vmatrix} g_1 & g_2 & \cdots & g_m \\ g_{21} & g_{22} & \cdots & g_{2m} \\ \cdot & \cdot & \cdots & \cdot \\ g_{m1} & g_{m2} & \cdots & g_{mm} \end{vmatrix} = 1 \tag{12}$$

if and only if $g_1, \cdots, g_m$ form a primitive set.

The g.c.d. of $g_1, \cdots, g_m$ is a factor of the first row of δ and divides δ. Thus $\delta = 1$ implies that $g_1, \cdots, g_m$ is primitive. Conversely let $g_1, \cdots, g_m$ be primitive. At least one $g_i \neq 0$ and we may assume that $g_1 \neq 0$. If $g_2 = \cdots = g_m = 0$ we have $g_1 h_1 = 1$ by (11) and take $g_{22} = h_1$, $g_{ij} = 0$ for $i \neq j$, all other $g_{ii} = 1$. Then $\delta = g_1 h_1 = 1$. If $g_2 \neq 0$ but $g_3 = \cdots = g_m = 0$ then $1 = g_1 h_1 + g_2 h_2$ by (11) and

$$\delta = \begin{vmatrix} g_1 & g_2 & 0 & 0 & \cdots & 0 \\ -h_2 & h_1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

[5] We are now using the notations of Section 1.

[6] For the case $m = 3$ see L. E. Dickson, *Studies in the Theory of Numbers*, p. 11. The result is undoubtedly well known in general but our proof is so simple that the author believes it a valuable addition to the paper. It actually seems to give a reason for the truth of the result.

Finally let g_1, g_2, g_3 be all non-zero so that $m > 2$ and the g.c.d. of $g_3, \cdots, g_m$ has the form

$$k = g_3k_3 + \cdots + g_mk_m \qquad (k_i \text{ in } \mathfrak{J}). \tag{13}$$

The integers g_1, g_2, k now form a primitive set. We let r be the product of all prime factors of g_1 not dividing g_2 and see that every prime factor of g_1 divides either rk but not g_2, or g_2 but not rk. Hence

$$s = g_2 + rk \tag{14}$$

is prime to g_1 and

$$g_1t - st_1 = 1 \qquad (t, t_1 \text{ in } \mathfrak{J}). \tag{15}$$

The determinant

$$\delta = \begin{vmatrix} g_1 & g_2 & g_3 & g_4 & \cdots & g_m \\ t_1 & t & 0 & 0 & \cdots & 0 \\ 0 & -k_3r & 1 & 0 & \cdots & 0 \\ 0 & -k_4r & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & -k_mr & 0 & 0 & \cdots & 1 \end{vmatrix} \tag{16}$$

is unaltered in value if we add the multiple rk_i of its i^{th} column to its second column for $i = 3, \cdots, m$. Since $s = g_2 + r(g_3k_3 + \cdots + g_mk_m)$ we have

$$\delta = \begin{vmatrix} g_1 & k & g_3 & g_4 & \cdots & g_m \\ t_1 & t & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

We have proved our lemma. The explicit form given by (13)–(16) for the case where at least three of our integers are not zero may itself be of some interest.

4. **Moduls over $\mathfrak{J}$.** Let $\mu_1, \cdots, \mu_m$ be linearly independent numbers of an algebraic field $\mathfrak{F}$ over $\mathfrak{R}$. Then the set

$$\mathfrak{M} = (\mu_1, \cdots, \mu_m)$$

consisting of all quantities

$$x_1\mu_1 + \cdots + x_m\mu_m \qquad (x_i \text{ in } \mathfrak{J}) \tag{17}$$

is an additive abelian group or modul which we shall call a modul over $\mathfrak{J}$ with basis $\mu_1, \cdots, \mu_m$.

Suppose also that $\mathfrak{M} = (\nu_1, \cdots, \nu_n)$, that is the ν_i form another basis of $\mathfrak{M}$ over $\mathfrak{I}$. Then

$$\nu_i = \sum_{j=1}^{m} g_{ij}\mu_j, \qquad \mu_j = \sum_{k=1}^{m} h_{jk}\nu_k \qquad (i, j = 1, \cdots, m),$$

with g_{ij} and h_{jk} in the domain $\mathfrak{I}$ of all rational integers. The product GH of the integral matrices $G = (g_{ij})$, $H = (h_{jk})$ is then the identity matrix I of m rows. For the ν_i are linearly independent in $\mathfrak{R}$. Then

$$|I| = |G| \cdot |H| = 1,$$

so that $\delta = |G|$ is an integer of $\mathfrak{I}$ dividing 1, and $\delta = \pm 1$. Conversely if $|G| = \pm 1$ the matrix $H = G^{-1}$ has elements in $\mathfrak{I}$ and $\mathfrak{M} = (\nu_1, \cdots, \nu_m)$. We have proved the well known theorem that *a set of m numbers $\nu_1, \cdots, \nu_m$ of a modul $\mathfrak{M} = (\mu_1, \cdots, \mu_m)$ over $\mathfrak{I}$ forms a basis of $\mathfrak{M}$ over $\mathfrak{I}$ if and only if the determinant of the coefficients of the expressions of the ν_i in terms of μ_j is ± 1.*

DEFINITION. A number

$$\phi = g_1\mu_1 + \cdots + g_m\mu_m \neq 0 \tag{18}$$

of $\mathfrak{M}$ is called a *basal element* of $\mathfrak{M}$ if $\mathfrak{M}$ has a basis with ϕ one of the basal numbers.

We apply our above criterion together with Lemma 2 and obtain

LEMMA 3. *A number ϕ of the form* (18) *is a basal element of $\mathfrak{M}$ if and only if $g_1, \cdots, g_m$ form a primitive set.*

Notice that in Lemma 2 we obtain $\delta = -1$ by interchanging two rows and that if $m = 1$ in Lemma 3 the result is trivial.

We shall require the generalization

LEMMA 4. *Let $\mathfrak{M} = (\mu_1, \cdots, \mu_m)$ and ϕ be given by* (18). *Then ϕ is an element of a basis*

$$\mu_1, \cdots, \mu_r, \phi, \phi_{r+2}, \cdots, \phi_m \tag{19}$$

for some $\phi_{r+2}, \cdots, \phi_m$ in $\mathfrak{M}$ if and only if the set of integers $g_{r+1}, \cdots, g_m$ of (18) *is primitive.*

For if $\mathfrak{M}$ has a basis (19) the determinant

$$\delta = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ g_1 & g_2 & \cdots & g_r & g_{r+1} & \cdots & g_m \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ g_{m,1} & g_{m,2} & \cdots & g_{m,r} & g_{m,r+1} & \cdots & g_{m,m} \end{vmatrix} \tag{20}$$

must be ± 1. The r-rowed Laplace expansion of δ gives

$$\delta = \begin{vmatrix} g_{r+1} & \cdots & g_m \\ \cdot & \cdots & \cdot \\ g_{m,r+1} & \cdots & g_{m,m} \end{vmatrix} = \pm 1 \tag{21}$$

so that the set $g_{r+1}, \cdots, g_m$ is primitive. Conversely if $g_{r+1}, \cdots, g_m$ is a primitive set we make δ of (21) equal to ± 1, and have δ of (20) equal to ± 1 for arbitrary $g_1, \cdots, g_r, g_{i1} \cdots, g_i$. Then the corresponding set (19) forms basis of $\mathfrak{M}$ over $\mathcal{I}$.

5. **Algebraic integers.** An algebraic number θ is an *algebraic integer* if it satisfies some $g(x) = x^r + b_1 x^{r-1} + \cdots + b_r = 0$ with integral b_i. The Gauss Lemma implies that θ is an algebraic integer if and only if the minimum function $f(x)$ of θ has integral coefficients. It also implies that a rational number is an algebraic integer if and only if it is an ordinary (rational) integer. Every polynomial in algebraic integers with coefficients in $\mathcal{I}$ is an algebraic integer. Thus the set $\mathcal{J}$ of all integers of $\mathcal{F}$ is what we shall call an integral domain containing $\mathcal{I}$.

The conjugates of an algebraic integer are also algebraic integers. Then so is

$$T(\phi) = \phi + \phi^{(2)} + \cdots + \phi^{(n)}. \tag{22}$$

Since $T(\phi)$ is rational it is an integer when ϕ is an integer. This so-called *trace function* is important and has the properties

$$T(1) = n, \qquad T(a\phi + b\psi) = aT(\phi) + bT(\psi) \tag{23}$$

for every rational a, b and every ϕ, ψ of $\mathcal{F}$. Here n is the degree of $\mathcal{F}$ over $\mathcal{R}$. Notice that $T(\theta) = a_1$ in (1) and that $T(\theta) = 0$ is equivalent to the taking of $f(x)$ to be what is called a reduced n-ic.

6. **Integral domains.** A modul $\mathfrak{M}$ over $\mathcal{I}$ is called an integral domain over $\mathcal{I}$ if the product of any two numbers of $\mathfrak{M}$ is in $\mathfrak{M}$. Thus $\mathfrak{M}$ is an integral domain if and only if there exist rational integers g_{ijk} such that

$$\mu_i \mu_j = \sum_{k=1}^{m} g_{ijk} \mu_k \qquad (i, j = 1, \cdots, m) \tag{24}$$

where we are assuming that $\mathfrak{M} = (\mu_1, \cdots, \mu_m)$. It is well known[7] that the set $\mathcal{J}$ of all algebraic integers of $\mathcal{F} = \mathcal{R}(\theta)$ is an integral domain

$$(\omega_1, \cdots, \omega_n) \tag{25}$$

over $\mathcal{I}$ and our interest is in the problem of determining the ω_i in terms of θ and its minimum function $f(x)$ of (1). We shall consider other integral domains however and shall prove

[7] Cf. L. E. Dickson, *Algebren und ihre Zahlentheorie*, s. 144.

LEMMA 5. *Let $\mathfrak{M} \leqq \mathfrak{F}$ be an integral domain over $\mathfrak{I}$. Then $\mathfrak{M} \leqq \mathfrak{J}$, that is the elements of $\mathfrak{M}$ are algebraic integers.*

For if ϕ is in $\mathfrak{M}$ it has the form (18). But $\mathfrak{M}$ is an integral domain over $\mathfrak{I}$ and there exist integers x_{ij} in $\mathfrak{I}$ such that

$$\phi\mu_i + \sum_{j=1}^{n} x_{ij}\mu_j = 0 \qquad (i = 1, \cdots, m). \tag{26}$$

The m linear homogeneous equations (26) in the μ_k have solutions μ_k not all zero and the customary theorem of the elementary theory of equations states that the determinant

$$h(\phi) = \begin{vmatrix} \phi + x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & \phi + x_{22} & \cdots & x_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ x_{m1} & x_{m2} & \cdots & \phi + x_{mn} \end{vmatrix} = 0.$$

But $h(\phi)$ is a polynomial in ϕ with coefficients in $\mathfrak{I}$ and leading coefficient unity. Hence ϕ is in $\mathfrak{J}$.

We shall apply Lemma 5 as follows. Let $\mathfrak{M}$ be an integral domain of order n over $\mathfrak{I}$ and let $\mathfrak{M}$ be maximal. Then $\mathfrak{M} \leqq \mathfrak{J}$ implies that $\mathfrak{M} = \mathfrak{J}$ and our basis of $\mathfrak{M}$ is an integral basis of $\mathfrak{F}$.

7. **The θ-bases of $\mathfrak{R}(\theta)$.** There are certain important special integral bases of $\mathfrak{F}$. Assume that θ is an algebraic integer with the property that $A\theta$ is integral for rational A if and only if A is integral.

It has then been shown[3] that the integers of $\mathfrak{F} = \mathfrak{R}(\theta)$ of degree n over $\mathfrak{F}$ have a basis

$$\omega_i = \frac{e_{i1} + e_{i2}\theta + \cdots + e_{i\,i-1}\theta^{i-2} + \theta^{i-1}}{E_i} \quad (i = 1, \cdots, n-1). \tag{27}$$

The numbers e_{ij}, E_i are rational integers and E_i divides the discriminant $d(\theta)$. Moreover

$$E_i \equiv 0(E_{i-1}) \qquad (i = 2, \cdots, n). \tag{28}$$

Since $\omega_1 = E_1^{-1}$ is integral we have

$$\omega_1 = 1. \tag{29}$$

The numbers $1, \theta, \cdots, \theta^{n-1}$ are linearly independent in the field $\mathfrak{R}$ and form a basis of $\mathfrak{R}(\theta)$ over $\mathfrak{R}$. This is also evidently true of the ω_i and their form immediately implies the

LEMMA 6. *Let ϕ be an algebraic integer of the form*

$$\phi = g^{-1}(g_1 + g_2\theta + \cdots + g_r\theta^{r-1}) \quad (g, g_i \text{ in } \mathfrak{I},\ r \leqq n-1).$$

Then ϕ is a rational integral linear combination of $\omega_1, \cdots, \omega_r$ of (27).

We apply Lemma 6 to $1, \theta, \cdots, \theta^{i-2}$ which are thus rational integral linear combinations of $\omega_1, \cdots, \omega_{i-1}$. Hence

$$\frac{(f_{i1} + f_{i2}\theta + \cdots + f_{i\,i-1}\theta^{i-2})E_i}{E_i} = \sum_{j=1}^{i-1} x_{ij}\omega_j,$$

for x_{ij} in $\mathcal{I}$. Define

$$\omega_{k0} = \omega_k, \qquad \omega_{i0} = \omega_i + \sum_{j=1}^{i-1} x_{ij}\omega_j \qquad (k \neq 1, k = 1, \cdots, n).$$

Then the ω_{k0} have the form (27) for $k = 1, \cdots, n$ and the determinant of the coefficients of the expressions is unity. But then the ω_{k0} are a basis (27) of $\mathfrak{R}(\theta)$ with

$$e_{ij}^0 = e_{ij} + f_{ij}E_i. \tag{30}$$

This implies that the e_{ij} are not unique but may be replaced by any integers congruent to them modulo E_i. Note then that if $E_i = 1$ we may take the $e_{ij} = 0$, $\omega_i = \theta^{i-1}$. This is a desirable transformation when possible and we make the

DEFINITION. A basis (27) of the integers of $\mathfrak{R}(\theta)$ will be called a *θ-basis* if $E_i = 1$ implies that $\omega_i = \theta^{i-1}$.

8. **The discriminant of $\mathcal{F}$.** The discriminant $d_{\mathcal{F}}$ of a field $\mathcal{F} = \mathfrak{R}(\theta)$ is the square of the determinant of the matrix of an integral basis of $\mathfrak{R}(\theta)$ and their conjugates given by

$$(\omega_i^{(j)}) \qquad (i, j = 1, \cdots, n).$$

The number $d_{\mathcal{F}}$ is an integral invariant of $\mathcal{F}$, that is $d_{\mathcal{F}}$ is unaltered by the replacement of $\omega_1, \cdots, \omega_n$ by any other integral basis of $\mathcal{F}$. We shall close this chapter, our preliminary discussion of algebraic numbers, by noticing the connection between $d_{\mathcal{F}}$, $d(\theta)$ and (27).

Write

$$\omega_i = \sum_{j=1}^{n} g_{ij}\theta^{j-1} \qquad (i = 1, \cdots, n)$$

and see that

$$(\omega_i^{(j)}) = (g_{ij})V_\theta \tag{31}$$

where V_θ is the Vandermonde matrix

$$V_\theta = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \theta & \theta^{(2)} & \cdots & \theta^{(n)} \\ \cdot & \cdot & \cdots & \cdot \\ \theta^{n-1} & (\theta^{(2)})^{n-1} & \cdots & (\theta^{(n)})^{n-1} \end{pmatrix}. \tag{32}$$

Then the discriminant $d(\theta) = |V_\theta|^2$ and

$$d_{\mathfrak{F}} = |g_{ij}|^2 d(\theta).$$

But $g_{ij} = 0$ for $j > i$ and $g_{ii} = E_i^{-1}$. Hence

$$d_{\mathfrak{F}} = \left(\prod_{i=2}^{n} E_i\right)^{-2} d(\theta). \tag{33}$$

This is an important restriction on the E_i when combined with (28).

Chapter II. Fields of Degree n over $\mathfrak{R}$

1. **Fundamental polynomials.** Our first normalization of the defining equation (1) of our algebraic number field $\mathfrak{F}$ of degree n over $\mathfrak{R}$ will be devised so as to satisfy the conditions in terms of which we gave the integral basis (27). We shall accomplish this by means of a linear transformation on θ and shall make θ have zero trace. Our results will be given in terms of polynomials defined by the

Definition. A polynomial $f(x)$ will be called a *fundamental polynomial* if

$$f(x) = x^n + a_2 x^{n-2} + \cdots + a_n \tag{34}$$

has rational integral coefficients, is irreducible in $\mathfrak{R}$, and is such that the simultaneous congruences

$$a_i \equiv 0(p^i) \qquad (i = 1, \cdots, n) \tag{35}$$

have no integral solution p other than $p = \pm 1$.

2. **Linear transformations.** The transformation

$$x = u + vy \qquad (u, v \neq 0 \text{ in } \mathfrak{R})$$

which replaces an indeterminate x by an indeterminate y is called a non-singular rational linear transformation on x. The set of all such transformations forms a group, so that every operation accomplished by a sequence of such transformations may also be accomplished by a single one.

Apply (34) to

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_n \qquad (a_i \text{ in } \mathfrak{R}). \tag{36}$$

We obtain the unique polynomial

$$g(y) = v^{-n} f(x) = y^n + A_1 y^{n-1} + \cdots + A_n. \tag{37}$$

The A_i are given by

$$A_i = A_i(u, v) = \frac{f_{n-i}(u)}{v^i} \qquad (i = 0, 1, \cdots, n), \tag{38}$$

where

$$f_j(x) = \frac{f^{(j)}(x)}{j!} \qquad (j = 0, \cdots, n+1), \tag{39}$$

and $f^{(j)}(x)$ is the j^{th} derivative of $f(x)$. Thus $f_0(x) = f(x)$, $f_{n+1}(x) = 0$ and the coefficients of $f_j(x)$ are always rational integral multiples of the coefficients a_i of $f(x)$. Notice that

$$A_i(0, v) = \frac{f_{n-1}(0)}{v^i} = \frac{a_i}{v^i} \qquad (i = 1, \cdots, n). \tag{40}$$

We compute $A_1(-a_1 n^{-1}, 1) = f_{n-1}(-a_1 n^{-1}) = 0$. Assume now that $a_1 = 0$. Write $a_i = e^{-1} c_i$ where e is a common denominator of the rational a_i and the c_i are integers. Then the transformation $x = e^{-1}y$ gives $A_i = e^i a_i$ all integral. Hence assume the a_i all integral and let p be the largest positive integer for which

$$a_i = p^i b_i,$$

with integral b_i. Then $x = py$ replaces the a_i by $A_i = p^{-i} a_i = b_i$. If $b_i = q^i c_i$ for an integer q then we may assume $q > 0$ and have $a_i = (pq)^i b_i$ whence $pq \geqq p$ must equal p, $q = 1$. We have proved that any $f(x)$ of (36) may be transformed into $f(x)$ of (34), (35). We now apply this result to algebraic number fields and prove

THEOREM 1. *Every algebraic number field $\mathcal{F}$ of degree n over $\mathcal{R}$ is generated by a root θ of a fundamental polynomial. Then θ is an algebraic integer of zero trace and such that if A is a rational number and $A\theta$ is integral then A is a rational integer.*

For, every $\mathcal{F} = \mathcal{R}(\theta)$ where θ is a root of an irreducible $f(x)$ of (36). Then also $\mathcal{F} = \mathcal{R}(\phi)$ for any rational u and $v \neq 0$ and $\theta = u + v\phi$. The minimum function of ϕ has degree n and is $g(y)$ of (37). Hence $g(y)$ is irreducible and our above proof states that u and v may be chosen so that $g(y)$ is a fundamental polynomial. Replace θ by ϕ and we have our normalized generation. Hence we may assume with no loss that $f(x)$ is fundamental. Write $A = ab^{-1}$ as a fraction in lowest terms so that a and $b > 0$ are relatively prime integers. The quantity $A\theta$ has (37) as minimum equation with $A_i = a_i(b^{-1}a)^i$. But $A\theta$ is integral if and only if the A_i are all integral and thus since a is prime to b, if and only if $a_i \equiv 0(b^i)$. By (35) we have $b = 1$, $A = a$ is an integer.

3. **Linear functions of θ.** The simplest type of an integral basis of $\mathcal{R}(\theta)$ would be one where $\omega_i = \theta^{i-1}$. In general no such basis exists and we must be satisfied with a more complicated one. However we may hope to obtain a basis of the type described in the

DEFINITION. A generation $\mathcal{R}(\theta)$ of a field $\mathcal{F}$ of degree n over $\mathcal{R}$ is called *basal* if θ is a root of a fundamental polynomial and $\omega_2 = \theta$ in any θ-basis of $\mathcal{F}$.

We shall use the notations of (38), (39) and prove

THEOREM 2. *Let $\mathcal{F} = \mathcal{R}(\theta)$ where θ is a root of a fundamental polynomial. Then θ provides a basal generation of $\mathcal{F}$ over $\mathcal{R}$ if and only if the congruences*

$$f_{n-i}(c_j) \equiv 0(p_j^i) \qquad (i = 0, \cdots, n) \tag{41}$$

have no simultaneous solution c_j for every prime divisor p_j of n.

For by (27) we have $\omega_2 = E^{-1}(c + \theta)$ with integral c, E and positive E. The minimum equation of ω_2 is given by (37) with

$$A_i = f_{n-i}(c)E^i. \tag{42}$$

Suppose first that (41) have no simultaneous solution. Then c is prime to E since otherwise $c = pc_0$, $E = pE_0$, with p an integer greater than unity, $-c_0 + E_0\omega_2 = \theta p^{-1}$ is an integer contrary to Theorem 1. The trace of ω_2 is $E^{-1}[nc + T(\theta)] = E^{-1}nc$ and is an integer. Hence E divides n. Since ω_2 is an integer the A_i of (42) are integral contrary to the non-solvability of (41) if $E > 1$. Hence $E = 1$, $\omega_2 = \theta$.

Conversely let $\omega_2 = \theta$. Then if p divides n and c satisfies $f_{n-i}(c) \equiv 0(p^i)$ our proof above shows that $p^{-1}(c + \theta)$ is an algebraic integer. Hence Lemma 6 implies that

$$p^{-1}(c + \theta) = g + h\theta \qquad (g, h \text{ in } \mathfrak{J})$$

which is impossible since p^{-1} is not integral.

Theorem 2 is applied in the discussion of polynomials $f(x)$ of (34) for which (41) have no simultaneous solution. It is not necessary to study (41) for all prime factors p_j of n when $d(\theta)$ is known. We may in fact prove the added restriction

THEOREM 3. *Let p be a prime factor of n and $f_{n-i}(c) \equiv 0(p^i)$ for $i = 0, \cdots, n$. Then $p^{n(n-1)}$ divides $d(\theta)$.*

For our hypothesis implies that $\rho = p^{-1}(c + \theta)$ is integral. Then the powers of ρ are all in $\mathfrak{J}$. The binomial theorem states that

$$\rho^i = k_{i0} + k_{i1}\theta + \cdots + \theta^i p^{-i} \qquad (i = 0, \cdots, n - 1)$$

for rational k_{ij}, and the Vandermonde matrix for ρ is

$$V_\rho = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ k_{10} & p^{-1} & 0 & \cdots & 0 \\ k_{20} & k_{21} & p^{-2} & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ k_{n-1,0} & k_{n-1,1} & k_{n-1,2} & \cdots & p^{-(n-1)} \end{pmatrix} V_\theta. \tag{43}$$

Hence

$$d(\rho) = |V_\rho|^2 = p^{-h}d(\theta)$$

must be an integer where $h = 2(1 + 2 + \cdots + n - 1) = n(n - 1)$.

Theorems 2 and 3 give a criterion by means of which we may determine all polynomials (34) defining algebraic number fields $\mathfrak{R}(\theta)$ with the property $\omega_2 = \theta$ in a θ-basis. We shall later obtain explicit results for the cases $n = 3, 4, 5$.

4. Existence of a θ-basis with $\omega_2 = \theta$. The study of the integral bases of quadratic number fields is complete and the results of course are very simple. We shall henceforth assume that $n > 2$.

Let $\omega_1 = 1, \omega_2, \cdots, \omega_n$ be a θ-basis of $\mathfrak{R}(\theta)$ and write

$$t_i = T(\omega_i) \qquad (i = 1, \cdots, n)$$

for the traces of the basal units ω_i. By our definition we may write

$$\omega_2 = \frac{-c + \theta}{E}, \qquad \phi_{jk} = \omega_2^{(j)} - \omega_2^{(k)} = \frac{\theta^{(j)} - \theta^{(k)}}{E} \neq 0$$

for $j \neq k$. If $c = 0$ then $\frac{\theta}{E}$ is an integer and $E = 1, \omega_2 = \theta$ as desired. Hence let $c \neq 0$ so that

$$t_2 = -ncE^{-1} \neq 0.$$

We shall prove

LEMMA 7. *There exists a primitive set of integers $g_2, \cdots, g_n$ such that $\rho = g_2\omega_2 + \cdots + g_n\omega_n$ has the properties*

$$T(\rho) \equiv 0(n), \qquad d(\rho) \neq 0.$$

If $t_3 + \cdots + t_n = 0$ we put $g_3 = \cdots = g_n = 1$, $g_2 = ng$ that $g_2, \cdots, g_n$ are a primitive set with $T(\rho) = ng_2t \equiv 0(n)$. Let $t = t_3 + \cdots + t_n \neq 0$ and suppose that

$$t_2 = s_2h, \qquad t = sh$$

where h is the greatest common divisor of t_2 and t. Then $g_2 = ns_2g + s$ and $g_3 = \cdots = g_n = -s_2$ form a primitive set for every integer g and $T(\rho) = (ns_2g + s)s_2h - s_2sh = n(s_2^2gh) \equiv 0(n)$ as desired. We now notice that

$$\rho^{(j)} - \rho^{(k)} = g_2\phi_{jk} + \psi_{jk}$$

where ψ_{jk} is a linear combination of $g_3, \cdots, g_n$ with algebraic number coefficients. Then

$$d(\rho) = \prod_{j<k} (g_2\phi_{jk} + \psi_{jk})^2$$

is a polynomial of degree 2ν in g_2 whose leading coefficient is the rational number

$$r = \prod_{j<k} \phi_{jk}^2 = \prod_{j<k}\left(\frac{\theta^{(j)} - \theta^{(k)}}{E}\right)^2 = \frac{d(\theta)}{E^{2\nu}} \neq 0,$$

We thus write $d(\rho) = rg_2^{2\nu} + \cdots$ and have either $d(\rho) = r(ng)^{2\nu} + \cdots$ or $d(\rho) = r(ns_2g + s)^{2\nu} + \cdots$. In either case the coefficient of $g^{2\nu}$ is not zero and $d(\rho) \neq 0$ for a sufficiently large integer g. This proves Lemma 7.

The statement $d(\rho) \neq 0$ implies that $\mathfrak{R}(\theta) = \mathfrak{R}(\rho)$ by Lemma 1. By (9) we have $\mathfrak{R}(\theta) = \mathfrak{R}(\phi)$ for $\phi = g_1 + \rho$, g_1 any rational integer. Since $T(\rho) = nv$ we take $g_1 = -v$ and obtain $T(-v + \rho) = -nv + nv = 0$. Since

$$\phi = g_1 + \rho = g_1\omega_1 + \cdots + g_n\omega_n$$

with $g_2, \cdots, g_n$ a primitive set we may apply Lemma 4 to show that there exists an integral basis of $\mathfrak{R}(\theta)$ with $\omega_{10} = 1$, $\omega_{20} = \phi$. Replace θ by ϕ and

assume then that $\mathfrak{R}(\theta)$ has an integral basis with $\omega_1 = 1$, $\omega_2 = \theta$. Let now $\mu_1, \mu_2, \cdots, \mu_n$ be any θ-basis (27) of $\mathfrak{R}(\theta)$. Then $\mu_2 = (g + \theta)E_2^{-1}$ and has the form $h_1\omega_1 + \cdots + h_n\omega_n$. But $\omega_1 = 1, \omega_2 = \theta, \cdots, \omega_n$ are linearly independent in $\mathfrak{R}$ and we must have $h_3 = \cdots = h_n = 0$, $(g + \theta)E_2^{-1} = h_1 + h_2\theta$, $E_2^{-1} = h_2\theta$ is an integer, $E_2 > 0$, so that $E_2 = 1$. By our definition $\mu_2 = \theta$ and we have proved

THEOREM 4. *Every algebraic number field $\mathfrak{F}$ of degree $n > 2$ over $\mathfrak{R}$ has a basal generation.*

5. **Two simultaneous congruences.** Consider the congruences

$$f(x) \equiv 0(E^2), \qquad f'(x) \equiv 0(E) \tag{44}$$

where both x and E are to be determined. Here $f(x)$ is any polynomial in x with integral coefficients and $f'(x)$ is the derivative of $f(x)$. We shall prove

THEOREM 5. *Let E_0 be the largest positive integer for which* (44) *have a simultaneous solution x. Then every E for which a simultaneous x exists divides E_0.*

If E is any integer of (44) and E_1 divides E then (44) are evidently satisfied for E_1 by the same x satisfying these congruences for E. We are actually proving a partial converse and as is usual in such cases, employ the Chinese remainder theorem. Let then E_1 and E_2 be relatively prime values of E in (44) with corresponding x_1, x_2. We may find an $x \equiv x_1(E_1^2)$, $x \equiv x_2(E_2^2)$ and then have $f(x) \equiv f(x_1) \equiv 0(E_1^2)$, $f(x) \equiv f(x_2) \equiv 0(E_2^2)$, and similarly for $f'(x)$ modulo E_1, E_2. The fact that E_1 and E_2 are relatively prime implies (44) with $E = E_1E_2$. Let D be the discriminant of $f(x)$, and $p_1, \cdots, p_r$ the distinct prime factors of D. Define e_i as the largest integer for which

$$f(x_i) \equiv 0(p_i^{2e_i}), \qquad f'(x_i) \equiv 0(p_i^{e_i}) \qquad (i = 1, \cdots, r)$$

have a simultaneous solution. Then our above result implies that (44) have a solution x for $E = E_0$ given by

$$E_0 = p_1^{e_1} \cdots p_r^{e_r}. \tag{45}$$

We now take any E of (44) and have $f(x) = f'(x) \equiv 0(p)$ for every prime factor p of E, that is[8] p divides D. Hence $E = p_1^{f_1} \cdots p_r^{f_r}$ and (44) implies that $f(x) \equiv 0(p_i^{2f_i})$, $f'(x) \equiv 0(p_i^{f_i})$ so that $f_i \leqq e_i$, E divides E_0 of (45), E_0 is the largest integer for which (44) have a simultaneous solution.

It would be interesting to consider further problems arising in the study of (44). For example it can be shown that when $f(x)$ is an irreducible cubic any two solutions of (44) modulo E_0 are congruent. We shall not require the generalization of this result to degree r however and it is not a part of our problem. We now show how these congruences arise.

---

[8] This is due to the well known theorem that $f(x) = 0$ and $f'(x) = 0$ have a common root if and only if $D = 0$. We are applying it in the field of residue classes modulo p.

6. **A type of modul $\mathfrak{M}$ over $\mathcal{J}$.** The integral domain $\mathcal{J}$ of all integers of $\mathcal{F} = \mathcal{R}(\theta)$ has a θ-basis (27). In our later work we shall assume that θ provides a basal generation of $\mathcal{F}$ but let us temporarily assume, for greater generality, that θ is a root of a fundamental polynomial

$$f(x) = x^n + a_2x^{n-2} + \cdots + a_n \qquad (a_i \text{ in } \mathcal{J}). \tag{46}$$

The work of finding (27) will then be materially assisted by considering moduls

$$\mathfrak{M} = (1, \theta, \cdots, \theta^{n-2}, \Omega) \tag{47}$$

over $\mathcal{J}$, where

$$\Omega = \frac{e_1 + e_2\theta + \cdots + e_{n-2}\theta^{n-3} + e\theta^{n-2} + \theta^{n-1}}{E} \tag{48}$$

with $e_1, \cdots, e_{n-2}, e, E$ in $\mathcal{J}$ and $E > 0$. Notice that $\mathfrak{M}$ consists of all rational integral linear combinations of $1, \theta, \cdots, \theta^{n-2}, \Omega$. Since $\mathfrak{M}$ contains $\theta^{n-1} = E\Omega - (e_1 + \cdots + e\theta^{n-2})$ it contains every rational integral polynomial in θ. We may then prove

LEMMA 8. *The modul $\mathfrak{M}$ of* (47), (48) *is an integral domain over $\mathcal{J}$ if and only if $\theta\Omega$ and Ω^2 are in $\mathfrak{M}$.*

For the products of the basal elements of $\mathfrak{M}$ are either powers of θ and hence in $\mathfrak{M}$, or Ω^2 assumed to be in $\mathfrak{M}$, or products $\theta^i\Omega$. If we assume that

$$\theta\Omega = g_1 + g_2\theta + \cdots + g_n\Omega \qquad (g_i \text{ in } \mathcal{J}) \tag{49}$$

is in $\mathfrak{M}$ then $\theta^2\Omega = g_1\theta + \cdots + g_{n-2}\theta^{n-1} + g_n\theta^n + g_n(g_1 + \cdots + g_n\Omega)$ is in $\mathfrak{M}$. An evident induction shows that the $\theta^i\Omega$ are all in $\mathfrak{M}$ as desired.

We wish to find necessary and sufficient conditions that a modul $\mathfrak{M}$ of (47), (48) be an integral domain. In order to be sure that such integral domains actually exist we shall prove

THEOREM 6. *Let $\mathfrak{M} = (\Omega_1, \cdots, \Omega_n)$ where $\Omega_i = E_i\omega_i$ $(i = 1, \cdots, n-1)$, $\Omega_n = E_{n-1}\omega_n$, and the ω_i, E_i are given by a θ-basis* (27) *of $\mathcal{R}(\theta)$. Then $\mathfrak{M} = (1, \theta, \cdots, \Omega_n)$ is an integral domain of* (47), (48).

For by (28) we have $E_{n-1} = D_iE_i$ for $i = 1, \cdots, n-2$ and $E_n = EE_{n-1}$. Hence $\Omega_i = e_{i_1} + \cdots + \theta^{i-1}$ and it is trivial to show that $\mathfrak{M}$ has a basis (47), (48) with $\Omega = \Omega_n$. Since $\mathcal{J}$ is an integral domain with $\omega_1, \cdots, \omega_n$ as basis over $\mathcal{J}$ we have $\omega_i\omega_j = \sum_{k=1}^n g_{ijk}\omega_k$ and g_{ijk} in $\mathcal{J}$. Hence

$$\Omega^2 = (E_{n-1}\omega_n)(E_{n-1}\omega_n) = \sum_{k=1}^n g_{nnk}E_{n-1}\omega_k = \sum_{k=1}^n (g_{nnk}D_k)\Omega_k$$

is in $\mathfrak{M}$ where of course $D_{n-1} = D_n = 1$. Now $\theta\omega_n = \sum_{k=1}^n h_k\omega_k$ is in $\mathcal{J}$ and $\theta\Omega = \theta(E_{n-1}\omega_n) = \sum_{k=1}^n h_kD_k\Omega_k$ is also in $\mathfrak{M}$. By Lemma 8 the modul $\mathfrak{M}$ is an integral domain over $\mathcal{J}$.

The hypothesis that $\mathfrak{M}$ is an integral domain enables us to normalize Ω. It is evident that if $\mathfrak{M}$ is given by (47) then $\mathfrak{M} = (1, \theta, \cdots, \theta^{n-2}, \Omega_0)$ for any Ω_0 which is the sum of Ω and a rational integral linear combination of $1, \theta, \cdots, \theta^{n-2}$. Take

$$\theta\Omega_0 = h_1 + h_2\theta + \cdots + h_{n-2}\theta^{n-3} + \Omega \qquad (h_i \text{ in } \mathcal{J}) \tag{50}$$

and use (49) to compute

$$
\begin{aligned}
\theta\Omega_0 &= h_1\theta + \cdots + h_{n-2}\theta^{n-2} + g_1 + \cdots + g_n\Omega \\
&= g_n\Omega_0 - g_n(h_1 + \cdots + h_{n-2}\theta^{n-3}) \\
&\quad + g_1 + \cdots + g_{n-1}\theta^{n-2} + h_1\theta + \cdots + h_{n-2}\theta^{n-2}.
\end{aligned} \tag{51}
$$

Then

$$\theta\Omega_0 = q + r\Omega_0 \qquad (q,\ r \text{ in } \mathfrak{I}), \tag{52}$$

if and only if

$$
\begin{aligned}
g_2 + h_1 - g_nh_2 = g_3 + h_2 - g_nh_3 = \cdots &= g_{n-2} + h_{n-3} - g_nh_{n-2} \\
&= g_{n-1} + h_{n-2} = 0.
\end{aligned} \tag{53}
$$

The i^{th} one of the equations (53) determines h_i uniquely in terms of h_{i+1}, g_n, g_{i+1} for $i = 1, \cdots, n-3$. The last determines $h_{n-2} = -g_{n-1}$. Thus (52) is true for uniquely determined h_i and we have the necessity condition

LEMMA 9. *If $\mathfrak{M}$ of (47), (48) is an integral domain the quantity (48) may be so chosen that $\theta\Omega = q + r\Omega$ with q, r in $\mathfrak{I}$.*

We now prove

THEOREM 7. *A modul $\mathfrak{M}$ of (47), (48) has the property $\theta\Omega = q + r\Omega$ with q, r in $\mathfrak{I}$ if and only if*

$$r = e,\quad qE = f(e),\quad e_{n-i} = e^i + a_2e^{i-2} + \cdots + a_i \quad (i = 2, \cdots, n-1). \tag{54}$$

When (54) is satisfied the modul $\mathfrak{M}$ is an integral domain over $\mathfrak{R}$ if and only if

$$f(e) \equiv 0(E^2), \qquad f'(e) \equiv 0(E). \tag{55}$$

For the powers $1, \theta, \cdots, \theta^{n-1}$ are linearly independent in $\mathcal{R}$ and

$$\theta\Omega = \frac{e_1\theta + \cdots + e_{n-2}\theta^{n-2} + e\theta^{n-1} - (a_2\theta^{n-2} + \cdots + a_n)}{E} = q + r\Omega$$

if and only if $r = e$, the last equations of (54) hold, and $-a_n = qE + ee_1$, $e^n + a_2e^{n-2} + \cdots + a_n = f(e) = qE$ as desired. We have of course obtained this result by equating coefficients of θ^{n-1}, the powers θ^i for $i = 1, \cdots, n-2$, and finally the constant terms respectively. We next assume (54) and see that $\theta\Omega = q + e\Omega$. We shall compute Ω^2 by a rather interesting device. Notice that

$$\theta^2\Omega = q\theta + e(q + e\Omega) = qe + q\theta + e^2\Omega.$$

Now if $\Omega^2 = g_1 + \cdots + g_n\Omega$ for g_i in $\mathfrak{I}$ we have $(\theta - e)\Omega = q$ and thus

$$
\begin{aligned}
q^2 &= (\theta^2 - 2e\theta + e^2)(g_1 + g_2\theta + \cdots + g_{n-1}\theta^{n-2} + g_n\Omega) \\
&= e^2\Phi_1 - 2e\Phi_2 + \Phi_3 + (g_{n-2} - 2eg_{n-1})\theta^{n-1} + g_{n-1}\theta^n + e^2g_n\Omega \\
&\quad - 2eg_n(q + e\Omega) + g_n(qe + q\theta + e^2\Omega),
\end{aligned} \tag{56}
$$

where

$$
\begin{cases}
\Phi_1 = g_1 + g_2\theta + \cdots + g_{n-1}\theta^{n-2}, \qquad \Phi_2 = g_1\theta + \cdots + g_{n-2}\theta^{n-2}, \\
\Phi_3 = g_1\theta^2 + \cdots + g_{n-3}\theta^{n-2}.
\end{cases} \tag{57}
$$

Since θ satisfies $f(\theta) = 0$ and Ω is given by (48) we have

(58) $\theta^n = -(a_2\theta^{n-2} + \cdots + a_n), \quad \theta^{n-1} = E\Omega - (e_1 + e_2\theta + \cdots + e\theta^{n-2}).$

The quantities $1, \theta, \cdots, \theta^{n-2}, \Omega$ are linearly independent in $\mathfrak{R}$ and the total coefficient of Ω in (56) must vanish after we make the replacements (58). This coefficient is

$$(g_{n-2} + 2eg_{n-1})E + e^2g_n - 2e^2g_n + e^2g_n = 0$$

if and only if

$$g_{n-2} = 2eg_{n-1}, \tag{59}$$

since $E \neq 0$. Hence the coefficient of θ^{n-1} in (56) is zero, equation (56) becomes

(60) $q^2 = e^2\Phi_1 - 2e\Phi_2 + \Phi_3 - g_{n-1}(a_2\theta^{n-2} + \cdots + a_n) - eq_nq + g_nq\theta.$

If $i + 1 = n - k$ then $k = n - i - 1$ and the coefficient of x^{i+1} in $f(x)$ is $a_k = a_{n-i-1}$. Thus the coefficient of θ^{i+1} in (54) is

$$e^2g_{i+2} - 2eg_{i+1} - g_{n-1}a_{n-i-1} + g_i = 0 \qquad (i = 1, \cdots, n-3), \tag{61}$$

while from 1 and θ we have

(62) $q^2 = e^2g_1 - g_{n-1}a_n + eqg_n, \qquad e^2g_2 - 2eg_1 + g_nq - q_{n-1}a_{n-1} = 0.$

Take $i = n - 3$ in (61), apply (59) and obtain

$$g_{n-3} = -e^2g_{n-1} + 2e(2eg_{n-1}) + a_{n-i-1}g_{n-1} = (3e^2 + a_2)g_{n-1}.$$

Similarly

$$\begin{aligned} g_{n-4} = -e^2g_{n-2} + 2eg_{n-3} + a_3g_{n-1} &= -2e^3g_{n-1} + 2e(3e^2 + a_2)g_{n-1} + a_3g_{n-1} \\ &= (4e^3 + 2ea_2 + a_3)g_{n-1}. \end{aligned}$$

This shows that the formula

$$g_i = [(n-i)e^{n-i-1} + (n-i-2)e^{n-i-3}a_2 + \cdots + a_{n-i-1}]g_{n-1} \qquad (i = 1, \cdots, n-3) \tag{63}$$

holds for $i = n - 3, n - 4$. Let it hold for values $i = j + 1, j + 2$ and obtain

$$\begin{aligned} g_j = &-e^2[(n-j-2)e^{n-j-3} + \cdots + (n-j-k-2)e^{n-j-k-3}a_k \\ &+ \cdots + a_{n-j-3}]g_{n-1} + 2e[(n-j-1)e^{n-j-2} \\ &+ \cdots + (n-j-k-1)e^{n-j-k-2}a_k + \cdots + a_{n-j-2}]g_{n-1} \\ &+ a_{n-j-1}g_{n-1} \\ = &[(2n - 2j - 2 - n + j + 2)e^{n-j-1} \\ &+ \cdots + (2n - 2j - 2k - 2 - n + j + k + 2)e^{n-j-k-1}a_k \\ &+ \cdots + a_{n-j-1}]g_{n-1} \\ = &[(n-j)e^{n-j-1} + \cdots + (n-j-k)e^{n-j-k-1}a_k \\ &+ \cdots + a_{n-j-1}]g_{n-1} \end{aligned}$$

which is (63) for $i = j$. This completes our induction and proves that (63), (59) are equivalent to (61), (59). However (63) uniquely determines the g_i for $i = 1, \cdots, n - 3$ in terms of e and g_{n-1} and we now need only satisfy (62). Notice that (63) states that

$$
(64)\qquad \begin{aligned} g_1 &= [(n-1)e^{n-2} + (n-3)e^{n-4}a_2 + \cdots + a_{n-2}]g_{n-1} \\ g_2 &= [(n-2)e^{n-3} + (n-4)e^{n-5}a_2 + \cdots + a_{n-3}]g_{n-1} \end{aligned}
$$

and we must use these formulas in (62). We substitute them and obtain

$$
(65)\qquad \begin{aligned} &[(n-2)e^{n-1} + (n-4)e^{n-3}a_2 \\ &+ \cdots + a_{n-3}e^2 - 2(n-1)e^{n-1} - 2(n-3)e^{n-3}a_2 \\ &- \cdots - 2ea_{n-2} - a_{n-1}]g_{n-1} \\ &= -[ne^{n-1} + (n-2)e^{n-3}a_2 + \cdots + a_{n-1}]g_{n-1} = -f'(e)g_{n-1} = qg_n \end{aligned}
$$

from the second equation of (62). The first becomes

$$
(66)\qquad \begin{aligned} q^2 &= [(n-1)e^n + (n-3)e^{n-2}a_2 \\ &+ \cdots + a_{n-2}e^2 - a_n - ne^n - (n-2)e^{n-2}a_2 \\ &- \cdots - a_{n-1}e]g_{n-1} = -f(e)g_{n-1} \end{aligned}
$$

by the use of (65). The quantity $q = f(e)E^{-1}$ in (64) and (67) gives $[f(e)]^2E^{-2} = -f(e)g_{n-1}$. We know that $f(e) \neq 0$ since $f(x)$ is irreducible in $\mathfrak{R}$. Hence

$$
(67)\qquad g_{n-1} + \frac{-f(e)}{E^2}
$$

that is

$$
(68)\qquad f(e) \equiv 0\ (E^2).
$$

Replacing q and g_{n-1} by their values in (65) we obtain $f'(e)f(e)E^{-2} = f(e)E^{-1}g_n$ so that

$$
(69)\qquad g_n = [f'(e)]E^{-1}, \qquad f'(e) \equiv 0\ (E).
$$

We have proved that if $\mathfrak{M}$ is an integral domain the conditions (55) hold. Conversely if (55) hold we define the g_i by (67), (68), (63) and have Ω^2 in $\mathfrak{M}$. But (54) imply that $\theta\Omega$ is in $\mathfrak{M}$ and $\mathfrak{M}$ is an integral domain by Lemma 8.

An integral domain $\mathfrak{M}$ with E the largest integer for which (55) have a solution is certainly a maximal integral domain (47), (48). For suppose that $\mathfrak{M} \leqq \mathfrak{M}_0$ where $\mathfrak{M}_0 = (1, \theta, \cdots, \theta^{n-2}, \Omega_0)$ and

$$
(70)\qquad \Omega_0 = E_0^{-1}(d_1 + \cdots + d\theta^{n-2} + \theta^{n-1}).
$$

Since E_0 must satisfy $f(d) \equiv 0\ (E_0)^2$, $f'(d) \equiv 0\ (E_0)$ by Theorem 7 we have E_0 a divisor of E by Theorem 5. But $\mathfrak{M} \leqq \mathfrak{M}_0$,

$$
\Omega = h_1 + h_2\theta + \cdots + h_n\Omega_0 \quad \text{and} \quad E^{-1} = h_nE_0^{-1}, \quad E_0 = h_nE
$$

is divisible by E. Since E_0, E are both positive we have $E_0 = E$, $h_n = 1$ and $\Omega_0 = -h_1 - h_2\theta - \cdots - h_{n-1}\theta^{n-2} + \Omega$ is in $\mathfrak{M}$ so that $\mathfrak{M}_0 = \mathfrak{M}$, $\mathfrak{M}$ is maximal.

The above application of Theorem 5 may be utilized in finding a θ-basis of a cubic field as follows. We use a basal generation of $\mathfrak{R}(\theta)$ and determine the largest E and a solution e of (54). Then $1, \theta, \Omega$ of (54), (48) form a θ-basis of $\mathfrak{F}$. For in a basal generation $\mathfrak{J}$ is an integral domain (47) for $n = 3$ and contains $\mathfrak{M}$. Then $\mathfrak{M}$ is maximal, $\mathfrak{M} = \mathfrak{J}$. For fields of higher degree this is not true but the maximal integral domain $\mathfrak{M}$ will partially determine $\mathfrak{J}$.

Chapter III. Cubic Fields

1. **Basal generations.** Every cubic field $\mathfrak{F} = \mathfrak{R}(\theta)$ over $\mathfrak{R}$ is generated by a root θ of a fundamental polynomial

$$f(x) = x^3 + ax + b. \tag{71}$$

This now means that a and b are ordinary integers and $f(x) = 0$ has no rational integral root, b is not divisible by p^3 for any prime p such that p^2 divides a. We have shown that θ provides a basal generation of $\mathfrak{F}$ if and only if the congruences

$$x^3 + ax + b \equiv 0 \ (27), \qquad 3x^2 + a \equiv 0 \ (9) \tag{72}$$

have no simultaneous solution. We now prove

THEOREM 8. *Every cubic field is generated by a root θ of a fundamental polynomial such that*

$$4a^3 + 27b^2 \equiv 0 \ (3^6)$$

only if $b \equiv 0$ (3). Then $\mathfrak{F} = \mathfrak{R}(\theta)$ has an integral basis

$$1, \theta, \omega = \frac{e^2 + a + e\theta + \theta^2}{E}, \tag{73}$$

where E is the largest positive integer for which

$$e^2 + ae + b \equiv 0 \ (E^2), \qquad 3e^2 + a \equiv 0 \ (E), \tag{74}$$

have a simultaneous solution e.

We use Theorem 2 and thus seek all fundamental polynomials for which (72) have no solution. By Theorem 3 this is true if $4a^3 + 27b^2 \not\equiv 0$ (3^6). If $4a^3 + 27b^2 \equiv 0$ (3^6) and $b \equiv 0$ (3) then $a \equiv 0$ (9) and $(72)_2$ implies that $x \equiv 0$ (3), $(72)_1$ that $b \equiv -(x^3 + ax) \equiv 0$ (27) which is impossible. Hence we again have a basal generation. But if $b \not\equiv 0$ (3) and $4a^3 + 27b^2 \equiv 0$ (3^6) we have $a = 3a_1$, $4a_1^3 + b^2 \equiv 0$ (27). Since $b^2 \equiv 1$ (3) for any $b \not\equiv 0$ (3) we have $4a_1^3 \equiv -1$ (3), $(2a_1)^3 \equiv 1$ (3), $2a_1 \equiv 1$ (3). It follows that

$$(2a_1 - 1)^3 \equiv 8a_1^3 - 12a_1^2 + 6a_1 - 1 \equiv 0 \ (27),$$
$$3(2a_1 - 1)^2 \equiv 12a_1^2 - 12a_1 + 3 \equiv 0 \ (27).$$

Using $8a_1^3 \equiv -2b^2$ (27) we obtain $-2b^2 - 6a_1 + 2 \equiv 0$ (27), $b^2 + a - 1 \equiv 0$ (27), $(-b)^3 + a(-b) + b \equiv 0$ (27) which is $(72)_1$ for $x = -b$. The second congru-

ence $3b^2 + a_1 \equiv 0$ (9) is satisfied since $a_1 + b^2 \equiv 2 + 1 \equiv 0$ (3). Hence we have the first part of our theorem. But the integers of $\mathcal{F}$ now have a basis $1, \theta, \omega$ and we have seen that (74) must hold and $\Omega = \omega$ is given by (54).

In a later section we shall actually prove the existence of a basal generation of $\mathcal{F}$ with 3^6 not a divisor of $4a^3 + 27b^2$.

2. **The form E_ρ.** The integral domain $\mathcal{J}$ of all algebraic integers of $\mathcal{F}$ has a basis (73) over $\mathcal{I}$. We compute the multiplication table of $\mathcal{J}$. Write

$$(75) \qquad f(e) = e^3 + ae + b = BE^2, \qquad f'(e) = 3e^2 + a = CE.$$

Then (67), (69), (53), (73) imply that

$$(76) \quad \theta^2 = -(C + a) - e\theta + E\omega,\ \theta\omega = BE + e\omega,\ \omega^2 = -2eB - B\theta + C\omega.$$

Every basal generation of $\mathcal{F}$ over $\mathcal{R}$ is defined by an algebraic integer

$$(77) \qquad \rho = x + y\theta + z\omega$$

with integral x, y, z such that $T(\rho) = 0$, $\mathcal{F} = \mathcal{R}(\rho)$ and $\mathcal{J}$ has an integral basis $1, \rho, \phi$. But these quantities form a basis of $\mathcal{J}$ over $\mathcal{I}$ if and only if y and z are relatively prime. We may in fact prove

THEOREM 9. *The quantity $\rho = x + y\theta + z\omega$ defines a basal generation of $\mathcal{F}$ if and only if x, y, z are integers such that y and z are relatively prime,*

$$(78) \qquad 3x = -Cz.$$

For the trace of θ^2 is well known to be $-2a$. Since $T(\theta) = 0$ we have

$$(79) \qquad T(\omega) = E^{-1}[3(e^2 + a) - 2a] = C,$$

the integer of (75). Then $T(\rho) = 3x + Cz = 0$ if and only if (78) holds. Moreover ρ is in $\mathcal{R}$ if and only if $y = z = 0$ which is not possible when y and z are relatively prime. Hence $\mathcal{R}(\rho) = \mathcal{F}$ and ρ provides a basal generation.

The field $\mathcal{F} = \mathcal{R}(\rho)$ for every non-rational ρ of $\mathcal{F}$. Let $\rho = x + y\theta + z\omega$ where y and z are relatively prime integers, x is in $\mathcal{I}$, but let us not assume the trace condition (78). Then $\mathcal{F}$ has an integral basis with $1, \rho$ basal elements by Chapter I. We also see that then $A^{-1}(e + \rho)$ is integral for integral e, $A > 0$ only when $A = 1$. Then $\mathcal{F}$ has a ρ-basis

$$(80) \qquad 1, \rho, \phi = \frac{k_1 + k\rho + \rho^2}{E\rho} = x_1 + y_1\theta + z_1\omega$$

since ϕ is an algebraic integer. Thus x_1, y_1, z_1 are integers such that

$$(81) \qquad \begin{vmatrix} 1 & 0 & 0 \\ x & y & z \\ x_1 & y_1 & z_1 \end{vmatrix} = yz_1 - zy_1 = \pm 1.$$

If we replace ϕ by $-\phi$ we replace E_ρ by $-E_\rho$ and $yz_1 - zy_1$ by its negative. Let us do this when necessary and assume therefore that $yz_1 - zy_1 = 1$. We shall

now compute $\pm E_\rho$ and see that $E_\rho > 0$ implies that E_ρ is the absolute value of the expression we have obtained. We then have

THEOREM 10. *The integer E_ρ of (80) is the absolute value of*

$$Ey^3 + 3ey^2z + Cy^2z + Bz^3. \tag{82}$$

For proof we solve the equations for ρ, ϕ in terms of $1, \theta, \omega$ simultaneously for θ, ω and obtain

$$\theta = g_1 - z_1\rho - z\phi, \qquad \omega = g_2 - y_1\rho + y\phi \tag{83}$$

where the form of g_1, g_2 will not be needed. We compute

$$\rho^2 = x^2 + y^2\theta^2 + z^2\omega^2 + 2xy\theta + 2xz\omega + 2yz\theta\omega = H_1 + H_2\theta + H_3\omega. \tag{84}$$

By (76) we have

$$\begin{aligned} H_1 &= x^2 - y^2(e^2 + a) - 2eBz^2 + 2yzBE, \quad H_2 = 2xy - ey^2 - Bz^2, \\ H_3 &= 2xz + 2yze + z^2C + y^2E. \end{aligned} \tag{85}$$

Using (83) we obtain $\rho^2 = -k_1 - k\rho \pm E_\rho\phi$ where

$$\pm E\rho = H_3y - H_2z = 2xyz + 2y^2ze + yz^2C + y^3E - 2xyz + ey^2z + Bz^3 \tag{86}$$

and have (82) and our theorem. We shall use the form of H_1 later.

The cubic form (82) will be used to normalize our generation of $\mathcal{F}$ so as to give as simple an E in (73) as possible. Notice that for basal generations $\mathfrak{R}(\rho)$ we may take any relatively prime y and z in (82) except that if $C \not\equiv 0$ (3) we must take $z \equiv 0$ (3).

3. **The coefficient a_ρ.** A quantity ρ which gives a basal generation of $\mathcal{F}$ satisfies

$$g(y) = y^3 + a_\rho y + b_\rho = 0, \tag{87}$$

where $-2a_\rho = T(\rho^2)$. This latter property will be used to compute a_ρ. We have $\rho = x + y\theta + z\omega$ with $x = -\frac{1}{3}Cz$, $T(\omega) = C$ and we use (84), (85). Now

$$\begin{aligned} T(\rho^2) = 3H_1 + H_3C = 3x^2 - 3y^2(e^2 + a) - 6eBz^2 + 6yzBE + 2xzC \\ + 2yzeC + z^2C^2 + y^2EC. \end{aligned}$$

Since $3x^2 = \frac{1}{3}C^2z^2 - 2xyzC = -\frac{2}{3}z^2C^2$ we see that the total coefficient of z^2 in $T(\rho^2)$ is $\frac{2}{3}C^2 - 6eB$. Use $EC = 3e^2 + a$ and obtain $-2ay^2$. Finally the total coefficient of yz is $=6Be + 2eC$ and

$$a_\rho = -\tfrac{1}{2}T(\rho^2) = ay^2 + \left(3eB - \frac{C^2}{3}\right)z^2 - (3EB + eC)yz. \tag{88}$$

We shall use the quadratic form (88) to eliminate certain redundant cases from our final results.

4. **Generations with $4a^3 + 27b^2 \not\equiv 0$ (3^6).** We shall use the form G to prove the auxiliary

THEOREM 11. *Every cubic field is generated by a root θ of a fundamental polynomial $x^3 + ax + b$ with $4a^3 + 27b^2 \not\equiv 0$ (3^6). Such a generation is basal.*

For θ provides a basal generation of $\mathfrak{F}$ only when

$$(89) \qquad 3e^2 + a = EC, \qquad e^3 + ae + b = E^2B$$

have no simultaneous integral solutions e, E, B, C for $E \equiv 0$ (9). Hence for every basal generation we have

$$(90) \qquad E \not\equiv 0 \ (3) \text{ or } E \text{ exactly divisible by } 3.$$

Let the generation be basal but $4a^3 + 27b^2 \equiv 0$ (3^6). Then $b \equiv 0$ (3) by Theorem 8, $a \equiv 0$ (3), and evidently

$$(91) \qquad a \equiv b \equiv 0 \ (9).$$

We now satisfy $e^3 + ae + b \equiv 0$ (9), $3e^2 + a \equiv 0$ (3) by taking $e \equiv 0$ (3) and it follows that $E \equiv 0$ (3). However $b \not\equiv 0$ (27), $e^3 + ae \equiv 0$ (27) so that $B \not\equiv 0$ (3). We take $y = 0$, $z = 1$ in (84) and $x = -\frac{1}{3}C$ since $3e^2 + a \equiv 0$ (9), $C \equiv 0$ (3). Then $\rho = -\frac{1}{3}C + \omega$ provides a basal generation of $\mathfrak{F}$ with $G = B \not\equiv 0$ (3). If $4a_\rho^3 + 27b_\rho^2 \equiv 0$ (3^6) then $b_\rho \equiv 0$ (3) by Theorem 8 and the above proof shows that $a_\rho \equiv b_\rho \equiv 0$ (9), $G \equiv 0$ (3) a contradiction. We replace θ by ρ and have our desired generation

The quantity $d(\theta) = -4a^3 - 27b^2 = E^2 d_{\mathfrak{F}}$ where $d_{\mathfrak{F}}$ is the discriminant of $\mathfrak{F}$. We shall remove the powers of three from E, $4a^3 + 27b^2$, $d_{\mathfrak{F}}$ and in fact write $E = 3^\alpha E_3$, $d_{\mathfrak{F}} = 3^\beta d_3$. Then $-(4a^3 + 27b^2) = 3^{2\alpha+\beta}D$ where $D = E_2^3 d_3$ is prime to 3. Hence $d_3 \equiv D$ (3). For the square of E_3 prime to three is always $\equiv 1$ (3). We shall later use the usual analogous property $E_2^2 \equiv 1$ (8) for powers of 2 dividing $d(\theta)$.

5. **The powers of 3 dividing E and $d_{\mathfrak{F}}$.** We have already seen that in every basal generation $\mathfrak{F} = \mathfrak{R}(\theta)$ the integer $E \not\equiv 0$ (9). Hence E is prime to three or exactly divisible by 3. In certain basal generations of a field $\mathfrak{F}$ we may have $E \equiv 0$ (3) and choose a basal generation $\mathfrak{F} = \mathfrak{R}(\rho)$ with $E_\rho \not\equiv 0$ (3). However there are certain fields $\mathfrak{F}$ in which $E_\rho \equiv 0$ (3) for every basal ρ. We shall actually show that these are the fields in which $d_{\mathfrak{F}}$ is exactly divisible by 3. We state our final conditions on E and $d_{\mathfrak{F}}$ modulo 3 in

THEORER 12. *Let a and b range over all integers for which $f(x) = x^3 + ax + b$ is irreducible in $\mathfrak{R}$, a has no factor $p^2 > 1$ such that p^3 divides b, and one of the mutually exclusive conditions*

$$(92) \qquad a \not\equiv 0 \ (3)$$

$$(93) \qquad a \equiv 6 \ (9),\ b \text{ exactly divisible by } 3,$$

(94) $a \equiv 0\ (9),\ b \equiv 2, 4, 5, 7\ (9),$

(95) $a \equiv 6\ (9),\ b \equiv 1, 8\ (9),$

(96) $a \equiv 6\ (9),\ b \equiv 4, 5\ (9),$

(97) $a \equiv 0\ (9)$, *b exactly divisible by* 3,

(98) $a \equiv 6\ (9),\ b \equiv 0\ (9),$

(99) $a \equiv 0\ (9),\ b \equiv 1, 8\ (9),$

is satisfied. Then the corresponding cubic fields $\mathfrak{F} = \mathfrak{R}(\theta)$ defined by $f(\theta) = 0$ range over all cubic fields and fields defined by distinct conditions (92)–(99) *are inequivalent since the corresponding discriminants are given respectively by*

(92d) $d_{\mathfrak{F}} \not\equiv 0\ (3),$

(93d) $d_{\mathfrak{F}} = 3^3 d_3,\ d_3 \equiv 1\ (3),$

(94d) $d_{\mathfrak{F}} = 3^3 d_3,\ d_3 \equiv 2\ (3),$

(95d) $d_{\mathfrak{F}} = 3^4 d_3,\ d_3 \equiv 1\ (3),$

(96d) $d_{\mathfrak{F}} = 3^4 d_3,\ d_3 \equiv 2\ (3),$

(97d) $d_{\mathfrak{F}} = 3^5 d_3,\ d_3 \not\equiv 0\ (3),$

(98d) $d_{\mathfrak{F}} = 3 d_3,\ d_3 \equiv 1\ (3),$

(99d) $d_{\mathfrak{F}} = 3 d_3,\ d_3 \equiv 2\ (3).$

Every cubic field above has an integral basis (73) *with $E \not\equiv 0$ (3) in* (92)–(97) *so that there is no resulting condition on e modulo* 3. *But every field $\mathfrak{F}$ of* (98), (99) *has $e \equiv -b$ (3), E exactly divisible by* 3 *and the property that $E_\rho \equiv 0$ (3) for every basal generation $\mathfrak{F} = \mathfrak{R}(\rho)$.*

Our theorem above *does not* state that two distinct polynomials $x^3 + ax + b$ in the same class (92)–(99) are equivalent. It is not our purpose to obtain a canonical set of fields but rather to give conditions in terms of which E and $d_{\mathfrak{F}}$ have a simple form. We now pass to a proof of the above theorem by first proving

LEMMA 10. *Let $b \not\equiv 0$ (3). Then* (89) *have a solution e for which $E \equiv 0$ (3) if and only if $a = 3a_1$ for integral a_1, $b^2 + a - 1 \equiv 0$ (9), $e \equiv -b$ (3).*

The integer E of (89) is divisible by 3 if and only if $e^3 + ae + b \equiv 0$ (9), $3e^2 + a \equiv 0$ (3). If $b \not\equiv 0$ (3) then $e \not\equiv 0$ (3), $a = 3a_1$, $e^3 + ae + b \equiv e^3 + b \equiv 0$ (3). But $e^2 \equiv 1$ (3), $e \equiv -b$ (3), $e^3 \equiv -b^3$ (9), $ae \equiv -ab$ (9), $e^3 + ae + b \equiv -b^3 - ab + b \equiv -b(b^2 + a - 1) \equiv 0$ (9) if and only if $b^2 + a - 1 \equiv 0$ (9). Conversely $a = 3a_1$, $b^2 + a - 1 \equiv 0$ (9), $e \equiv -b$ (3) imply our congruences, and thus $E \equiv 0$ (3). This proves our lemma.

If $C \not\equiv 0\ (3)$ we have seen that $\rho = x + y\theta + z\omega$ defines a basal generation only if $z \equiv 0\ (3)$ so that, by (82), if $E \equiv 0\ (3)$ then $E_\rho \equiv 0\ (3)$. But $E \equiv 0\ (3)$ implies that $a = 3a_1$ for integral a_1. Also $b \not\equiv 0\ (3)$ implies that $e \equiv -b \not\equiv 0\ (3)$ while $b \equiv 0\ (3)$ implies that $e \equiv 0\ (3)$. We shall prove

LEMMA 11. *Let $C \not\equiv 0\ (3)$, $E \equiv 0\ (3)$. Then a generation of $\mathfrak{F}$ exists with $a \not\equiv 3\ (9)$.*

For let $a \equiv 3\ (9)$, $C \not\equiv 0\ (3)$, $E \equiv 0\ (3)$ so that we must take $z = 3z_1$, y prime to z in ρ. Then by (88)

$$a_\rho = 3[a_1 y^2 + 3(-EByz_1 + 3eBz_1^2) - C^2 z_1^2 - eCyz_1] = 3a_{\rho 1},$$

where $a_{\rho 1} \equiv a_1 y^2 - C^2 z_1^2 - eCyz_1 \equiv -eCyz_1$ if z_1 is prime to 3. Hence $a_\rho \equiv 0\ (9)$ if $e \equiv 0\ (3)$ while otherwise we may choose y so that $a_{\rho 1} \equiv -eCyz_1 \equiv 2\ (3)$. Then $a_\rho \equiv 6\ (9)$. In either case $E_\rho \equiv E \not\equiv 0\ (9)$, $E_\rho \equiv 0\ (3)$.

First let $a \equiv 0\ (9)$. If $b \equiv 0\ (3)$ then $e \equiv 0\ (3)$, $3e^2 \not\equiv a \equiv 0\ (9)$, $C \equiv 0\ (3)$ which is contrary to hypothesis. Hence $b \not\equiv 0\ (3)$ so that $E \equiv 0\ (3)$ if and only if $b^2 + a - 1 \equiv 0\ (9)$ by Lemma 10. Then $b^2 \equiv 1\ (9)$, $b \equiv 1, 8\ (9)$ and we have (99). Also $-4a^3 - 27b^2 = -27\,(4a_1^3 + b^2)$ where $4a_1^3 + b^2 \equiv b^2 \equiv 1\ (9)$. Then $d_3 \equiv -b^2 \equiv 2\ (3)$ by Section 4. Moreover E is exactly divisible by 3, $d(\theta)$ is exactly divisible by 3^3 so that $d_{\mathfrak{F}} = 3d_3$ and we have (99d). Next let $a \equiv 6\ (9)$ so that $3e^2 + a = 3(e^2 + a_1) \equiv 0\ (9)$ if and only if $e \not\equiv 0\ (3)$. Hence $e \equiv 0\ (3)$, $e^3 + ae + b \equiv b \equiv 0\ (9)$ and we have (98). We also obtain $-(4a^3 + 27b^2) = 27\,(-4a_1^3 + b^2)$, $d_3 \equiv -4a_1^3 \equiv -2 \equiv 1\ (3)$ and obtain (98d). Notice that in both cases $e \equiv -b\ (3)$ and that when our conditions are satisfied we do have E exactly divisible by 3, $C \not\equiv 0\ (3)$.

There remains the case where either $C \equiv 0\ (3)$ or $E \not\equiv 0\ (3)$. In the former case with $E \equiv 0\ (3)$ we must not have $B \equiv 0\ (3)$ since then $e^3 + ae + b \equiv 0\ (27)$, $3e^2 + a \equiv 0\ (9)$ which is impossible for a basal generation. Then $\rho = -\frac{1}{3}C + \omega$ gives the value $\pm B \not\equiv 0\ (3)$ for E_ρ so that we may assume with no loss of generality that in all remaining cases we have $E \not\equiv 0\ (3)$. We now prove a lemma analogous to Lemma 11.

LEMMA 12. *Let $E \not\equiv 0\ (3)$. Then there exists a basal generation of $\mathfrak{F}$ such that $E_\rho \not\equiv 0\ (3)$, $a_\rho \not\equiv 3\ (9)$.*

For let $a \equiv 3\ (9)$ and assume first that $b \not\equiv 0\ (3)$. The equations (89) impose no restrictions on e and we may if we like take $e \equiv 0\ (3)$, $C = 3C_0$, $C_0 = e^2 + a_1 \not\equiv 0\ (3)$. Then $\rho = -C_0 + \omega$ gives a basal generation with $a_\rho = 3eB - 3C_0^2 \equiv -3 \equiv 6\ (9)$, while $E_\rho = \pm B$. However $e^3 + ae + b \equiv b \not\equiv 0\ (3)$ so that $B \not\equiv 0\ (3)$ and $E_\rho \not\equiv 0\ (3)$ as desired. We next let $b \equiv 0\ (3)$. We again choose $e \equiv 0\ (3)$, $C = 3C_0$, $C_0 \not\equiv 0\ (3)$ and have $B \equiv 0\ (3)$ since $E \not\equiv 0\ (3)$, $e^3 + ae + b \equiv 0\ (3)$. Then if $\rho = -C_0 + \theta + \omega$ we have

$$E_\rho \equiv \pm(E + 3e + C + B) \equiv \pm E \not\equiv 0\ (3)$$

while $a_p = a + (3eB - 3C_0^2) - (3EB + 3eC_0) \equiv a - 3C_0^2 \equiv 3 - 3 \equiv 0\ (9)$. This proves our lemma.

If $a \not\equiv 0\ (3)$ then we have (92), (92d). Hence let $a = 3a_1$ so that, as we have seen, $a \equiv 0\ (9)$ or $a \equiv 6\ (9)$. If $a \equiv 6\ (9)$ and $b \equiv 0\ (3)$ then

$$e^3 + ae + b \equiv 0\ (9)\ 3e^2 + a \equiv 0\ (3)$$

if and only if $b \equiv 0\ (9)$, $e \equiv 0\ (3)$. Hence $E \not\equiv 0\ (3)$ if and only if $b \not\equiv 0\ (9)$, b is exactly divisible by three. Thus we have (93) and $-(4a^3 + 27b^2) = 27\,(-4a_1^3 - b^2)$ where $-(4a_1^3 + b^2) \equiv -2 \equiv 1\ (3)$ so that we obtain (93d). Next let $b \not\equiv 0\ (3)$ so that by Lemma 10 we must not have $b^2 \equiv 1 - a \equiv 4\ (9)$. But a square is congruent to 1, 4, 7 (9) so that $b^2 \equiv 1, 7\ (9)$. In the former case we have (95) and $-(4a_1^3 + b^2) \equiv -(32 + 1) \equiv 3\ (9)$ so that $d_{\mathcal{F}} = 3^4 d_3$, $d_3 \equiv 1\ (3)$ and we have (95d). In the latter case we have (96) and

$$-(4a_1^3 + b^2) \equiv -(32 + 7) \equiv -3 \equiv 6\ (9)$$

and we have (96d). Finally let $a \equiv 0\ (9)$. If $b \equiv 0\ (3)$ then again $E \not\equiv 0\ (3)$ if and only if $b \not\equiv 0\ (9)$ so we obtain (97). Moreover $-(4a^3 + 27b^2) = -27 \cdot 9\,(4 \cdot 3a_0^3 + b_0^2)$ if $a = 9a_0$, $b = 3b_0$. Hence $d_{\mathcal{F}} = 3^5 d_3$. We finally let $b \not\equiv 0\ (3)$ and by Lemma 10 have $b^2 \not\equiv 1\ (9)$. Then $b^2 \equiv 4, 7\ (9)$ and we have (94). We obtain (94d) from the fact that $b^2 \equiv 1\ (3)$, $-(4a^3 + 27b^2) \equiv -27\,(4 \cdot 27 + b^2)$, $d_3 \equiv -b^2 \equiv 2\ (3)$.

Theorem 12 gives a final set of conditions on our constants modulo 3. We now pass to a study of their values modulo 2.

6. **The divisors 2^k of E and $d_{\mathcal{F}}$.** The theory of our constants modulo 2 is quite analogous to that modulo 3. We shall actually prove the existence of certain fields with $d(\xi) \equiv 0\ (4)$ for every integer ξ of $\mathcal{F}$ but with $d_{\mathcal{F}}$ odd. These fields have been called cubic fields with a non-essential discriminant divisor. We state our results in

THEOREM 13. *Every cubic field $\mathcal{F}$ is generated by a root of a polynomial satisfying the conditions of Theorem* 12 *and such that one of the mutually exclusive conditions*

(100)	*b is odd.*	
(101)	$b \equiv 2\ (4)$,	$a \equiv 0\ (2)$,
(102)	$b \equiv 0\ (4)$,	$a \equiv 1\ (4)$,
(103)	$b \equiv 2\ (4)$,	$a \equiv 3\ (4)$,
(104)	$b \equiv (a + 1)\ (8)$,	$a \equiv 3\ (4)$.

Fields $\mathfrak{F}$ defined by distinct conditions (100)–(104) *are inequivalent since the respective discriminants have the form*

(100d) $d_{\mathfrak{F}}$ *odd,*

(101d) $d_{\mathfrak{F}} = 4d_2$, $d_2 \equiv 1\ (4)$,

(102d) $d_{\mathfrak{F}} = 4d_2$, $d_2 \equiv 3\ (4)$,

(103d) $d_{\mathfrak{F}} = 8d_2$, d_2 *odd,*

(104d) $d_{\mathfrak{F}}$ *odd,*

while in (100)–(103) *E is odd and there is no condition on e modulo* 2, *but in* (104) *the integer E is double an odd, e must be odd, and $d(\xi) \equiv 0$* (4) *for every integral ξ of $\mathfrak{F}$.*

For let θ satisfy the conditions of Theorem 12. Then $\rho = x + y\theta + z\omega$ also satisfies these conditions if $z \equiv 0\ (9)$, $x = -\frac{1}{3}zC$, $y \equiv 1\ (9)$. For it is quite clear that this choice leaves all the values of a, b invariant modulo 3. We may thus choose y and z in an arbitrary fashion modulo powers of 2 so long as our assumptions do not make both even.

Let $E \equiv 0\ (4)$. If $B \not\equiv 0\ (4)$ we make $E_\rho \not\equiv 0\ (4)$ in (82) by taking z odd and $y \equiv 0\ (4)$. If $B \equiv 0\ (4)$ and $C \equiv e \equiv 0\ (2)$ then $e^3 + ae + b = E^2B$ is divisible by 8, $3e^2 + a \equiv 0\ (4)$ so that $a \equiv 0\ (4)$, $b \equiv 0\ (8)$ which is contrary to hypothesis. Hence if either e or C is even we make E_ρ odd by taking $y \equiv z \equiv 1\ (2)$. Let then C and e both be odd and take y odd, $z \equiv C(2 - 3ey)\ (4)$ so that z is odd, $E_\rho = yz(Cz + 3ey) \equiv 2yz \equiv 2\ (4)$. This proves that there exists a generation $\mathfrak{R}(\theta)$ with the conditions of Theorem 12 satisfied and $E \not\equiv 0\ (4)$.

Let $E = 2E_2$ where E_2 is odd. Then either $B \equiv 0\ (2)$, $e \equiv C \equiv 1\ (2)$ or we may make E_ρ odd. Hence assume $B \equiv 0\ (2)$, $e \equiv C \equiv 1\ (2)$ so that $3e^2 + a \equiv 2\ (4)$, $a \equiv 3\ (4)$, $e^3 + ae + b \equiv 0\ (8)$. Then $e(a + 1) + b \equiv 0\ (8)$. Since $a + 1 \equiv 0\ (4)$ we have $e(a + 1) \equiv -(a + 1)\ (8)$, $b \equiv (a + 1)\ (8)$. This of course means that $b \equiv 0, 4\ (8)$ according as $a + 1$ is or is not divisible by eight. This gives (104) and (104d) comes from $4a^3 + 27b^2 = 4(a^3 + 27b_2^2)$, $b = 2b_2$, $b_2 \equiv 0\ (2)$ so that $a^3 + 27b_2^2 \not\equiv 0\ (2)$. Conversely let (104) hold. Then the only solutions of $3e^2 = a = EC$, $e^3 + ae + b = E^2B$ with E even are those with e odd, $3e^2 + a \equiv a + 3 \equiv 2\ (4)$ so that C is odd, $e^3 + ae + b \equiv e(a + 1) + b \equiv 0\ (8)$ so that B is even. But then $B = 2B_2$, $E = 2E_2$,

$$2E_2y^3 + 3ey^2z + Cy^2z^2 + 2B_2 \equiv 0\ (2)$$

so that every $E_\rho \equiv 0\ (2)$. But the discriminant of ρ is $E_\rho^2 d_{\mathfrak{F}} \equiv 0\ (4)$ while $d_{\mathfrak{F}}$ is odd as we have seen. Every integer $\xi = x + u(y\theta + z\omega)$ with relatively

prime y and z and if we put $\rho = y\theta + z\omega$ we have $d(\xi) = u^6 d(\rho) \equiv 0$ (4). This proves the last part of our theorem.

Let $E \not\equiv 0$ (2). If b is odd then $d(\theta) = -(4a^3 + 27b^2)$ is odd and so is $d_{\mathfrak{F}}$. This gives (100), (100d). Let $b = 2b_2$ and first assume that $b \equiv 2$ (4). If a is even then $3e^2 + a \equiv 0$ (2) only if e is even. Then $e^3 + ae + b \equiv b \not\equiv 0$ (4), so that E is odd as desired. This gives (101) and $-(4a^3 + 27b^2) = 4(-a^3 - 27b_2^2)$ where $a \equiv 0$ (2) implies that $d_{\mathfrak{F}} = 4d_2$, $d_2 \equiv -27b_2^2 \equiv -27 \equiv 1$ (4) and we have (101d). Let a be odd so that $a \equiv 3$ (4) or $a \equiv 1$ (4). In the latter case $e^3 + ae + b \equiv 2e + b \equiv 0$ (4) if e is odd, $3e^2 + a \equiv 0$ (2) so that E is even contrary to hypothesis. Hence $a \equiv 3$ (4) and $3e^2 + a \equiv 0$ (2) implies that e is odd, $e^3 + ae + b \equiv e(1 + a) + b \equiv b \equiv 2$ (4) so that E is odd and (103) holds. However $-(a^3 + 27b_2^2) \equiv -(a + 3) \equiv -6 \equiv 2$ (4) so that $d_{\mathfrak{F}} = 8d_2$ where d_2 is odd and this gives (103d). Finally let $b = 4b_2$, the only remaining case. If $a \equiv 3$ (4) we have $3e^2 + a \equiv 0$ (2) for odd e, $e^3 + ae + b \equiv e(a + 1) \equiv 0$ (4) and E is even, a contradiction. If $a \equiv 0$ (2) then $e \equiv 0$ (2) implies that E is even. Hence $a \equiv 1$ (4) whence conversely $3e^2 + a \equiv 0$ (2) only if e is odd, $e^3 + ae + b \equiv e(a + 1) \equiv 2$ (4) and E is odd. This gives (102), $-(a^3 + 27b_2^2 \cdot 4) \equiv -a^3 \equiv -a \equiv 3$ (4), and we have (102d). We have proved Theorem 13.

7. **The prime factors of E, $d_{\mathfrak{F}}$ which are prime to six.** The replacement of θ by $\rho = x + y\theta + z\omega$ with $y \equiv 1$ (72), $z \equiv 0$ (72), $x = -\frac{1}{3}Cz$ and y prime to z evidently does not alter the conditions of Theorems 12, 13. We may then choose the residues y and z modulo primes p neither 2 nor 3 in an arbitrary fashion so long as y and z are relatively prime. Let $p \neq 2, 3$ be a prime. If p divides e, E, C, B then $b \equiv 0$ (p^3), $a \equiv 0$ (p^2) which is impossible. Hence one coefficient of (82) is prime to p. If $E \not\equiv 0$ (p) we take y prime to p, $z \equiv 0$ (p) and have $E_\rho \not\equiv 0$ (p). We have a similar result when $B \not\equiv 0$ (p). If $E \equiv B \equiv 0$ (p) and either e or C is divisible by p we take y and z both prime to p and have $E_\rho \not\equiv 0$ (p). Finally let $E \equiv B \equiv 0$ (p), $\pm E_\rho \equiv yz(3ey + Cz)$ (p) with e and C prime to p. If $3e + C \not\equiv 0$ (p) we find relatively prime integers y and z congruent to unity modulo p and have $E_\rho \not\equiv 0$ (p). Otherwise $-C \equiv 3e$ (p), $y \equiv 1$ (p), $z \equiv 2$ (p) gives $\pm E_\rho \equiv 2(3e + 2C) \equiv 2C \not\equiv 0$ (p).

If g is any integer and $p_1, \cdots, p_t$ are the distance prime factors of g which are neither 2 nor 3 we apply the above result to each p_i in turn. Our congruences modulo p_i may be satisfied independently of those modulo $p_j \neq p_i$ by the Chinese remainder theorem. This gives

THEOREM 14. *Let $\mathfrak{F}$ be a cubic field and g be any integer prime to 6. Then $\mathfrak{F} = \mathfrak{R}(\theta)$ where θ satisfies the conditions of Theorems 12, 13 and the corresponding integer E is prime to g.*

Theorem 14 will be applied to simplify our determination of $d_{\mathfrak{F}}$. For we have

THEOREM 15. *Let $\mathfrak{F}$ be a cubic field with discriminant $d_{\mathfrak{F}}$ and write $d_{\mathfrak{F}} = 2^\gamma 3^\delta D_{\mathfrak{F}}$*

where $D_{\mathfrak{F}}$ is prime to 6. *Then $\mathfrak{F}$ has a basal generation $\mathfrak{R}(\theta)$ with a, b satisfying the conditions of Theorems* 12, 13 *and E prime to $D_{\mathfrak{F}}$.*

We next prove

LEMMA 13. *Let $\mathfrak{F} = \mathfrak{R}(\theta)$ where θ is as in Theorem* 15 *and $\theta^3 + a\theta + b = 0$. Then if p is a prime divisor of a, its square does not divide b.*

For let $a \equiv 0\ (p)$, $b \equiv 0\ (p^2)$. Then $e \equiv 0\ (p)$ implies that $3e^2 + a \equiv 0\ (p)$, $e^3 + ae + b \equiv 0\ (p)$. By Theorem 5 $E \equiv 0\ (p)$. If $E \equiv 0\ (p^2)$ then $e^3 + ae + b \equiv 0\ (p^4)$, $3e^2 + a \equiv 0\ (p^2)$ imply that $e \equiv 0\ (p)$, $a \equiv 0\ (p^2)$, $b \equiv 0\ (p^3)$ which is impossible. But $4a^3 + 27b^2 \equiv 0\ (p^3)$ so that $E^2 d_{\mathfrak{F}} \equiv 0\ (p^3)$, $d_{\mathfrak{F}} \equiv 0\ (p)$. This contradicts Theorem 15.

Let θ be as in Theorem 15 and G be the g.c.d. of a and b. Write

$$a = Ga_0, \qquad b = Gb_0$$

so that a_0 and b_0 are relatively prime. By Lemma 15 the integer G has no square factors and is prime to b_0. Hence

$$d(\theta) = -(4a^3 + 27b^2) = G^2H, \qquad H = -(4a_0^3G + 27b_0^2).$$

But H is prime to all prime factors of ab neither 2 nor 3. For if p divides G then p does not divide b_0 and hence H. If p divides a_0 then again p does not divide $27b_0^2$ and is prime to H. Finally if p divides b_0 it is prime to $4a_0^3G$ and hence to H.

We write

$$H = -4a_0^3G - 27b_0^2 = 2^\lambda 3^\mu Q^2 D$$

where D is prime to b and has no square factors. Finally write

$$E = 2^\gamma 3^\delta E_0$$

where E_0 is prime to 6. Then the author has shown[9] that the congruences $e^3 + ae + b \equiv 0\ (E_0^2)$, $3e^2 + a \equiv 0\ (E_0)$ are equivalent to

$$E_0 = Q, \qquad 2ae + b \equiv 0\ (Q).$$

We have proved that $D_{\mathfrak{F}} = G^2D$. Moreover $4a^3 + 27b^2$ is not exactly divisible by an odd power of a prime $p \neq 2, 3$. For G is prime to $H = Q^2D$ and $Q = E_0$ is prime to D which has no square factors.

Conversely let a and b be as in Theorems 12, 13 and such that $4a^3 + 27b^2$ is not exactly divisible by an odd power of a prime $p \neq 2, 3$. Then $a \equiv 0\ (p)$, $b \equiv 0\ (p^2)$, $p \neq 2, 3$ imply that $a = pa_0$, $b = p^2b_0$ with $a_0b_0 \not\equiv 0\ (p)$, $4a^3 + 27b^2 = p^3(4a_0^3 + 27pb_0^2)$ is exactly divisible by p^3, a contradiction. Moreover Q is then prime to D and we have proved our final result

---

[9] See pp. 562, 563 of the *Annals* paper cited in footnote 4. Notice that our present choice of θ makes $P = 1$ in that paper. Also identify our e with c there.

THEOREM 16. *Every cubic field $\mathfrak{F}$ has a generation $\mathfrak{F} = \mathfrak{R}(\theta)$, where a and b are integers satisfying the conditions of Theorems* 12, 13 *and such that* $4a^3 + 27b^2$ *is not exactly divisible by an odd power of any prime* $p \neq 2, 3$. *Then if G is the g.c.d. of a and b, it has no square factor, and we write*

$$-(4a^3 + 27b^2) = G^2Q^2 2^\lambda 3^\mu D, \tag{105}$$

where D has no square factors and is prime to ab. Then

$$E = 2^\alpha 3^\beta Q, \qquad d_{\mathfrak{F}} = G^2 2^{\lambda-2\alpha} 3^{\mu-2\beta} D \tag{106}$$

with β, α having the values 0, 1 *and determined as in Theorems* 12, 13 *respectively. Moreover all conditions on e of* (73) *are satisfied by taking*

$$e \equiv 1\ (2), \qquad e \equiv -b\ (3), \qquad 2ae + b \equiv 0\ (Q). \tag{107}$$

In closing we notice first that we have shown in Theorem 13 that $d_{\mathfrak{F}}$ may be exactly divisible by $2^0, 2^2, 2^3$. In Theorem 12 we showed that $d_{\mathfrak{F}}$ may be exactly divisible by $3^0, 3^1, 3^3, 3^4, 3^5$. If p is any prime not 2, 3 then Theorem 16 implies that $d_{\mathfrak{F}}$ is exactly divisible by p if p is prime to a and by p^2 if p divides both a and b.

Notice finally that it is possible to choose relatively prime integers a and b so that $4a^3 + 27b^2$ is exactly divisible by p^t for any prime $p \neq 2, 3$ and exponent $t > 0$. This implies that Q is quite arbitrary. For $4x + 27 \equiv p^t\ (p^{t+1})$ has a solution x_0 which is prime to p. Then $4x_0^3 + 27x_0^2 \equiv p^t x_0^2\ (p^{t+1})$. The values $a = x_0$, $b = x_0 + p^{t+1}$ make a prime to b and $4a^3 + 27b^2 \equiv p^t x_0\ (p^{t+1})$ so that $4a^3 + 27b^2$ is exactly divisible by p^t. Of course these values may give a reducible cubic. To obtain an irreducible cubic easily we choose $a = up^t + x_0$, $b = vp^t + x_0$ to be distinct positive primes of the infinitely many primes in the arithmetic progression $np^t + x_0$ with relatively prime coefficients. Then $x^3 + ax + b = 0$ is reducible if and only if it has integral roots. These divide b and must be $x = \pm 1, \pm b$. But then $x^3 + ax$ is even and cannot equal $-b$. We have proved the existence of a cubic field defined by $x^3 + ax + b = 0$ with Q divisible by p^t and have thus shown the condition in the first sentence of Theorem 16 is certainly non-trivial.

8. **The minimum of E.** The smallest possible value of E_ρ is the least positive value of the form

$$E(y, z) = Ey^3 + 3ey^2z + Cyz^2 + Bz^3$$

for relatively prime y and z. For evidently when this form has a negative value we make it positive by changing the signs of y and z. When $E(y, z) = 1$ the field $\mathfrak{F}$ has an integral basis

$$1, \rho, \rho^2 \tag{108}$$

and this is the simplest possible form of such a basis. It is clear that in the case (104) this is impossible since every E_ρ is even. In the cases (98), (99) this could be possible but not for $T(\rho) = 0$ since then $E_\rho \equiv 0$ (3). But even in more normal cases where there are no such reasons the form $E(y, z)$ can have a minimal value greater than unity and thus no integral basis $1, \rho, \rho^2$. We shall prove this result by giving the following concrete example.

Consider the polynomial $f(x) = x^3 - 7x + 14$. This cubic has $a = -7$, $b = 14$ and the only possible integral roots are $\pm 1, 2, 7, 14$. These are easily verified to be not roots and thus the cubic is irreducible. Now

$$-d(\theta) = 4a^3 + 27b^2 = 4(-7^3) + 27 \cdot 7^2 \cdot 4 = 2^4 \cdot 5 \cdot 7^2.$$

The congruences $e^3 - 7e + 14 \equiv 0$ (16), $3e^2 - 7 \equiv 0$ (4) have no solution. For otherwise e is odd, $e = 4n + 1$ or $4n + 3$. In the former case $e^3 \equiv 12n + 1$ (16), $e^3 - 7e + 14 \equiv 12n + 1 - 28n - 7 + 14 \equiv 8$ (16), while in the latter $e^3 \equiv 36n + 27 \equiv 4n + 11$,

$$e^3 - 7e + 14 \equiv 4n + 11 - 28n - 21 + 14 \equiv 4(1 - 6n) \not\equiv 0 \ (16).$$

However $e = 7$ gives

$$3e^2 + a = 3 \cdot 7^2 - 7 = 2 \cdot 70, \qquad e^3 + ae + b = 7^3 - 7^2 + 14 = 4 \cdot 77.$$

The g.c.d. of a and b is 7 and $d(\theta)$ is prime to 3. Thus $G = 1$ and

$$E = 2, \qquad d_{\mathfrak{J}} = -4 \cdot 5 \cdot 7^2.$$

We have not used Theorem 13 since we would have obtained a much more complicated but odd E if we had used the normalization of that theorem. However we now have

$$E = 2, \qquad C = 70, \qquad B = 77, \qquad e = 7 \tag{109}$$

so that

$$E(y, z) = 2y^3 + 21y^2z + 70yz^2 + 77z^3. \tag{110}$$

We shall show that 2 is the least E_ρ, that is $E(y, z) \neq 1$ for any integral y, z.

For if $E(y, z) = 1$ we have $E(y, z) \equiv 1$ (7). Then $2y^3 \equiv 1$ (7) and $4y^6 \equiv 2y^3 \equiv 1$ (7). But $y^6 \equiv 1$ (7) for any y prime to 7 by the Fermat theorem. Hence $4 \equiv 1$ (7) which is untrue.

Chapter IV. Basal Generations of Quartic and Quintic Fields

1. **Quartic fields.** The final result in our theory of basal generations of cubic fields was expressed in terms of the discriminant of our cubic. The discriminants of quartics and quintics are very complicated functions of the coefficients and we shall obtain simple but not analogous results in these cases. We first study irreducible quartics

$$x^4 + ax^2 + bx + c, \tag{111}$$

where a, b, c are integers such that if p is a prime and $a \equiv 0\ (p^2)$, $b \equiv 0\ (p^3)$, $c \not\equiv 0\ (p^4)$. We then wish the congruences

$$e^4 + ae^2 + be + c \equiv 0\ (16), \quad 4e^3 + 2ae + b \equiv 0\ (8), \quad 6e^2 + a \equiv 0\ (4)$$

to have no simultaneous solution e. Suppose first that they have a simultaneous solution. If e is even then $a \equiv 0\ (4)$, $b \equiv 0\ (8)$, $c \equiv 0\ (16)$, a contradiction. Hence e is odd and $a = 2a_1$, $3e^2 + a_1 \equiv 1 + a_1 \equiv 0\ (2)$ so that a_1 is odd. Then

$$4e^3 + 2ae \equiv 4e(e^2 + a_1) \equiv 4e(1 + a_1) \equiv 0\ (8)$$

so that $b \equiv 0\ (8)$, $b = 8b_2$. Now $e^2 \equiv 1\ (8)$ so that $e^4 \equiv 1\ (16)$, $ae^2 = 2a_1e^2 \equiv 2a_1 \equiv a\ (16)$ since $2(8n + 1) \equiv 2\ (16)$. Finally $be \equiv b\ (16)$ since $b \equiv 0\ (8)$ and e is odd. Our congruences thus imply that $a + b + c + 1 \equiv 0\ (16)$. The converse is trivial since all congruences are satisfied by $e = 1$. We have proved

THEOREM 17. *A quartic* (111) *defines a basal generation of a quartic field if and only if the conditions*

$$\text{(112)} \qquad a \text{ double an odd}, \qquad b \equiv 0\ (8), \qquad a + b + c + 1 \equiv 0\ (16)$$

are not simultaneously satisfied.

Our conditions are evidently of a much more simple nature than any involving the discriminant of (111). It is clear that our conditions for cubic fields could be stated easily in terms of the discriminant of $x^3 + ax + b$ principally because of the simple form of that quantity.

2. **Quintic fields.** A quintic field is generated by a root of the irreducible equation

$$\text{(113)} \qquad x^5 + ax^3 + bx^2 + cx + d = 0,$$

with rational integers $a, \cdots, d$ such that if p is a positive integer and $a \equiv 0\ (p^2)$, $b \equiv 0\ (p^3)$, $c \equiv 0\ (p^4)$, $d \equiv 0\ (p^5)$ then $p = 1$. The equation (113) provides a basal generation of $\mathfrak{R}(\theta)$ if and only if

$$\text{(114)} \qquad \begin{aligned} e^5 + ae^3 + be^2 + ce + d &\equiv 0\ (5^5) \\ 5e^4 + 3ae^2 + 2be + c &\equiv 0\ (5^4) \\ 10e^3 + 3ae + b &\equiv 0\ (5^3) \\ 10e^2 + a &\equiv 0\ (5^2) \end{aligned}$$

have no simultaneous solution e. We shall determine necessary and sufficient conditions for the simultaneous solvability of (114). The problem is a much more difficult one than in our previous cases both because of the larger number of congruences, the high degree of the first congruence, and the large moduli $5^5, 5^4, \cdots, 5$. We first prove

LEMMA 14. *Let* (114) *have a solution. Then* $a = 5g$, $b = 5h$, $c = 5k$ *and* (114) *are equivalent to the simultaneous congruences*

$$e^5 + 5(ge^3 + he^2 + ke) + d \equiv 0 \ (5^5), \tag{115}$$

$$e^4 + 3ge^2 + 2he + k \equiv 0 \ (5^3), \tag{116}$$

$$2e^3 + 3ge + h \equiv 0 \ (5^2), \tag{117}$$

$$2e^2 + g \equiv 0 \ (5), \tag{118}$$

for e, g *prime to* 5.

For $(114)_4$ implies that $a \equiv 0$ (5), $a = 5g$. Then $b \equiv 0$ (5) by $(114)_3$, $c \equiv 0$ (5) by $(114)_2$, and we have the equivalent congruences (115)–(118). If $g \equiv 0$ (5) then $e \equiv 0$ (5) by (118) and $h \equiv 0$ (5^2) by (117), $k \equiv 0$ (5^3) by (116), $d \equiv 0$ (5^5). Then $a \equiv 0$ (5^2), $b \equiv 0$ (5^3), $c \equiv 0$ (5^4), $d \equiv 0$ (5^5) which is impossible. Hence g is prime to 5. But then (118) implies that e is prime to 5.

We now prove

LEMMA 15. *The congruences* (117), (118) *are simultaneously satisfied if and only if*

$$2ge + h \equiv 0 \ (5), \qquad 2g^3 + h^2 \equiv 0 \ (25), \tag{119}$$

so that h *is prime to* 5.

For by (118) we have $2e^3 + ge \equiv 0$ (5) and subtracting from (117) we have $2ge + h \equiv 0$ (5). This is the first congruence of (119). We square it and obtain

$$4g^2e^2 + 4ghe + h^2 \equiv 0 \ (25) \tag{120}$$

from which $4g^2e^3 + 4ghe^2 + h^2e \equiv 0$ (25). Now $2g^2$ is prime to 5 and thus (117) is equivalent to its product by $2g^2$, that is

$$4g^2e^3 + 6g^3e + 2g^2h \equiv 0 \ (25). \tag{121}$$

Thus $(119)_1$, (120) imply that the congruence $4ghe^2 + (h^2 - 6g^3)e - 2g^2h \equiv 0$ (25) is equivalent to (117). But this is equivalent to

$$4g^2he^2 + g(h^2 - 6g^3)e - 2g^3h \equiv 0 \ (25) \tag{122}$$

while (120) gives $4g^2he^2 + 4gh^2e + h^3 \equiv 0$ (25). Hence $2ge + h \equiv 0$ (5) implies that (117) is equivalent to

$$(3gh^2 + 6g^4)e + h^3 + 2g^3h \equiv 0 \ (25),$$

that is to

$$(h^2 + 2g^3)(3ge + h) \equiv 0 \ (25). \tag{123}$$

Now $3ge + h \equiv 3ge - 2ge \equiv ge \not\equiv 0$ (5) so that (123) implies that $h^2 + 2g^3 \equiv 0$ (25) and we have (119). Conversely when (119) holds we see that (117) is satisfied by retracing the steps above. Moreover (118) is equivalent to $4g^2e^2 + 2g^3 \equiv 0$ (5) and by (120) to $-4ghe - h^2 + 2g^3 \equiv 0$ (5) by (119) to

$-2h(-h) - h^2 + 2g^3 \equiv 2g^3 + h^2 \equiv 0$ (5) which is true by $(119)_2$. This congruence and g prime to 5 imply that h is prime to 5 and we have proved our lemma.

Congruence (120) is obtained by squaring $(119)_1$. In a similar fashion we shall use the powers

$$(124)\quad (2ge + h)^3 = 8g^3e^3 + 12g^2he^2 + 6gh^2e + h^3 \equiv 0\ (5^3),$$

$$(125)\quad (2ge + h)^4 = 16g^4e^4 + 32g^3he^3 + 24g^2h^2e^2 + 8gh^3e + h^4 \equiv 0\ (5^4),$$

$$(126)\quad (2ge + h)^5 = 32g^5e^5 + 80g^4he^4 + 80g^3h^2e^3 + 40g^2h^3e^2 + 10gh^4e + h^5 \equiv 0\ (5^5),$$

and shall prove

LEMMA 16. *Let* (119) *be satisfied. Then* (116) *is true if and only if*

$$(127)\quad g^3 + 2h^2 \equiv 4gk\ (5^3),$$

so that k is prime to 5.

For (116) is equivalent to

$$(128)\quad 8g^3e^4 + 24g^4e^2 + 16g^3he + 8g^3k \equiv 0\ (5^3)$$

while (124) implies that

$$(129)\quad 8g^3e^4 + 12g^2he^3 + 6gh^2e^2 + h^3e \equiv 0\ (5^3).$$

Then (129), which is a consequence of $2ge + h \equiv 0$ (5), implies that (116) is equivalent to

$$(130)\quad 12g^2he^3 + 6g(h^2 - 4g^3)e^2 + h(h^2 - 16g^3)e - 8g^3k \equiv 0\ (5^3).$$

Since $2g$ is prime to 5 this is equivalent to

$$(131)\quad 24g^3he^3 + 12g^2(h^2 - 4g^3)e^2 + 2gh(h^2 - 16g^3)e - 16g^4k \equiv 0\ (5^3).$$

But (124) is a consequence of $2ge + h \equiv 0$ (5) and implies that

$$(132)\quad 24g^3he^3 + 36g^2h^2e^2 + 18gh^3e + 3h^4 \equiv 0\ (5^3)$$

by multiplication by $3h$. Hence (131) is equivalent to

$$(133)\quad 24g^2(h^2 + 2g^3)e^2 + 16gh(h^2 + 2g^3)e + 3h^4 + 16g^4k \equiv 0\ (5^3).$$

We use (120) to show that $24g^2e^2 + 16ghe \equiv -24ghe - 6h^2 + 16ghe \equiv -8ghe - 6h^2 \equiv -4h(-h) - 6h^2 \equiv -2h^2$ (5). Then $h^2 + 2g^3 \equiv 0$ (25) implies that $(24g^2e^2 + 16ghe)(h^2 + 2g^3) \equiv -2h^2(h^2 + 2g^3)$ (5^3), and (134) is equivalent to $-2h^4 - 4h^2g^3 + 3h^4 - 16g^4k \equiv 0$ (5^3). Hence (116) is equivalent to

$$(134)\quad h^4 \equiv 4h^2g^3 - 16g^4k\ (5^3).$$

But by (119) we have $h^4 + 4g^3h^2 + 4g^6 \equiv 0$ (5^4) so that (134) is equivalent to $-(4g^3h^2 + 4g^6) \equiv 4g^3h^2 - 16g^4k$ (5^3), $8g^3h^2 + 4g^6 - 16g^4k \equiv 0$ (5^3) which is true if and only if we have (127) since $4g^3$ is prime to 5. But $2h^2 \equiv -4g^3$, $g^3 + 2h^2 \equiv -3g^3 \equiv 4gk$ (5) and k is prime to 5.

Congruence (115) is equivalent to

$$(135) \qquad 32g^5e^5 + 160g^6e^3 + 160hg^5e^2 + 160kg^5e + 32g^5d \equiv 0 \ (5^5).$$

We use (126) and see that (115) is equivalent to

$$(136) \qquad \begin{aligned} 80g^4he^4 + 80g^3(h^2 - 2g^3)e^3 + 40g^2h(h^2 - 4g^3)e^2 + 10g(h^4 - 16g^4k)e \\ + h^5 - 32g^5d \equiv 0 \ (5^5). \end{aligned}$$

But (125) implies that

$$(137) \qquad 80g^4he^4 + 160g^3h^2e^3 + 120g^2h^3e^2 + 40gh^4e + 5h^5 \equiv 0 \ (5^5)$$

so that (115) is equivalent to

$$(138) \qquad \begin{aligned} 80g^3(h^2 + 2g^3)e^3 + 80g^2h(h^2 + 2g^3)e^2 + 10g(3h^4 + 16g^4k)e + 4h^5 \\ + 32g^5d \equiv 0 \ (5^5). \end{aligned}$$

Since $20(h^2 + 2g^3)g \equiv 0 \ (5^3)$ we have

$$(139) \quad 80g^3(h^2 + 2g^3)e^3 + 80g^2h(h^2 + 2g^3)e^2 + 20gh^2(h^2 + 2g^3)e \equiv 0 \ (5^5)$$

from (120). Then (115) is equivalent to

$$(140) \qquad 10g(h^4 - 4g^3h^2 + 16g^4k)e + 4h^5 + 32g^5d \equiv 0 \ (5^5).$$

By our proof of (127) we have (134) so that $5(h^4 - 4g^3h^2 + 16g^4k)(2ge + h) \equiv 0 \ (5^5)$. Hence (140) is equivalent to

$$(141) \qquad -5h(h^4 - 4g^3h^2 + 16g^4k) + 4h^5 + 32g^5d \equiv 0 \ (5^5).$$

Now $5(h^4 + 4g^3h^2 + 4g^6) \equiv 0 \ (5^5)$ so that

$$-5h(-8g^3h^2 - 4g^6 + 16g^4k) + 4h^5 + 32g^5d \equiv 0 \ (5^5)$$

and hence (141) and thus (115) is equivalent to

$$(142) \qquad 5hg^3(2h^2 + g^3 - 4gk) + h^5 + 8g^5d \equiv 0 \ (5^5).$$

Evidently (142) implies in particular that

$$(143) \qquad h^5 + 8g^5d \equiv 0 \ (5^4),$$

so that d is prime to 5.

Let $(119)_2$, (127), (142) hold for g, h, k prime to 5. Then $2ge + h \equiv 0 \ (5)$ determines e uniquely modulo 5 and can always be satisfied. But then we have shown that (115)–(118) hold and have proved the final result given by the

THEOREM 18. *An irreducible quintic* (113) *provides a basal generation of the corresponding field if one of a, b, c is not exactly divisible by* 5. *Otherwise* $a = 5g$, $b = 5h$, $c = 5k$ *for integers g, h, k prime to* 5, *and* (113) *provides a basal generation if and only if one of the congruences*

$$(144) \qquad \begin{aligned} 2g^3 + h^2 \equiv 0 \ (5^2), \qquad\qquad 4gk \equiv g^3 + 2h^2 \ (5^3), \\ 8g^5d \equiv 5hg^3(4gk - g^3 - 2h^2) - h^5 \ (5^5), \end{aligned}$$

is not satisfied.

The second and third congruences of (144) determine the form of k and d respectively with respect to h and g since $4g$ and $8g^5$ are prime to 5. The first of these congruences implies that $-2g$ is a quadratic residue of 25. Then

$$-2g \equiv 1, 4, 6, 9, 11, 14, 16, 19, 21, 24 \ (25),$$

so that multiplying by 12

$$g \equiv 12, 23, 22, 8, 7, 18, 17, 3, 2, 13 \ (25).$$

Thus $-2g \equiv u^2$ (25) for a unique integer $u = 1, 2, 3, 4, 6, 7, 8, 9, 11, 12$ and $-2g^3 \equiv (gu)^2 \equiv h^2(25)$ if and only if $h \equiv \pm gu$ (25). Hence $(144)_1$ may be thought of as determining both g and h in a partial way. This shows that the congruences (114) are independent and that we have given a non-redundant set of congruences whose simultaneous solution is necessary and sufficient for a generation to be non-basal.

THE UNIVERSITY OF CHICAGO.

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

SECOND SERIES, VOL. 38, NO. 1
JANUARY, 1937

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

ANNALS OF MATHEMATICS

Subscription price, $5 a volume (four numbers) in advance. Single copies $1.50. An arrangement has been made with the Mathematical Association of America in accordance with which an individual member of the Association may subscribe to the Annals of Mathematics at one-half the regular price. Subscriptions, orders for back numbers and changes of address should be sent to Mount Royal and Guilford Avenues, Baltimore, Md., or to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to The Annals of Mathematics, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten, with the exception of formulae, and must be in final form with all references filled in.

Authors receive gratis 100 reprints of each article, postage prepaid.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in United States of America

TABLE OF CONTENTS

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

S. LEFSCHETZ J. VON NEUMANN
F. BOHNENBLUST

WITH THE COÖPERATION OF THE
DEPARTMENT OF MATHEMATICS OF PRINCETON UNIVERSITY
AND
THE SCHOOL OF MATHEMATICS OF THE INSTITUTE
FOR ADVANCED STUDY
AND

SECOND SERIES, VOL. 38, NO. 2
APRIL, 1937

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

ANNALS OF MATHEMATICS

Subscription price, five dollars a volume (four numbers) in advance. Single copies one dollar and fifty cents. An arrangement has been made with the Mathematical Association of America by which an individual member of the Association may subscribe to the Annals of Mathematics at one-half the regular price. Subscriptions, orders for back numbers and changes of address should be sent to Mount Royal and Guilford Avenues, Baltimore, Maryland, or preferably to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to the Annals of Mathematics, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced. Footnotes should be numbered consecutively and typed (double-spaced) on a separate sheet; they will not be inserted in the text until page proofs are made.

Figures and diagrams should be drawn in India ink on separate sheets of paper and twice the size they are to be printed. The lettering is to be omitted from these drawings, but a sketch with complete lettering should be appended.

Authors ordinarily receive galley proofs only. Stylistic changes must be avoided in proof; excessive corrections will be charged to the authors. Authors receive gratis, postage prepaid, one hundred reprints of each article; joint authors receive fifty reprints each.

Suggestions to Authors

Footnotes should be reduced to a minimum or replaced whenever possible by a bibliography at the end of the paper; formulae in footnotes should be avoided. Authors are requested to keep in mind the typographical difficulties of complicated formulae. For example, "exp" may be used to avoid complicated exponents and a power to replace the troublesome radical sign. Symbols set one on top of another present a typographical difficulty which in most cases can easily be avoided by a stepwise arrangement.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in United States of America

9

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

SECOND SERIES, VOL. 38, NO. 3

JULY, 1937

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

ANNALS OF MATHEMATICS

Subscription price, five dollars a volume (four numbers) in advance. Single copies one dollar and fifty cents. An arrangement has been made with the Mathematical Association of America by which an individual member of the Association may subscribe to the Annals of Mathematics at one-half the regular price. Subscriptions, orders for back numbers and changes of address should be sent to Mount Royal and Guilford Avenues, Baltimore, Maryland, or preferably to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to the Annals of Mathematics, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced. Footnotes should be numbered consecutively and typed (double-spaced) on a separate sheet; they will not be inserted in the text until page proofs are made.

Figures and diagrams should be drawn in India ink on separate sheets of paper and twice the size they are to be printed. The lettering is to be omitted from these drawings, but a sketch with complete lettering should be appended.

Authors ordinarily receive galley proofs only. Stylistic changes must be avoided in proof; excessive corrections will be charged to the authors. Authors receive gratis, postage prepaid, one hundred reprints of each article; joint authors receive fifty reprints each.

Suggestions to Authors

Footnotes should be reduced to a minimum or replaced whenever possible by a bibliography at the end of the paper; formulae in footnotes should be avoided. Authors are requested to keep in mind the typographical difficulties of complicated formulae. For example, "exp" may be used to avoid complicated exponents and a power to replace the troublesome radical sign. Symbols set one on top of another present a typographical difficulty which in most cases can easily be avoided by a stepwise arrangement.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March, 3, 1879.

Made in United States of America

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

SECOND SERIES, VOL. 38, NO. 4

OCTOBER, 1937

PUBLISHED QUARTERLY
AT
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.
BY THE
PRINCETON UNIVERSITY PRESS
PRINCETON, N. J.

ANNALS OF MATHEMATICS

Subscription price, five dollars a volume (four numbers) in advance. Single copies one dollar and fifty cents. An arrangement has been made with the Mathematical Association of America by which an individual member of the Association may subscribe to the Annals of Mathematics at one-half the regular price. Subscriptions, orders for back numbers and changes of address should be sent to Mount Royal and Guilford Avenues, Baltimore, Maryland, or preferably to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to the Annals of Mathematics, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced. Footnotes should be numbered consecutively and typed (double-spaced) on a separate sheet; they will not be inserted in the text until page proofs are made.

Figures and diagrams should be drawn in India ink on separate sheets of paper and twice the size they are to be printed. The lettering is to be omitted from these drawings, but a sketch with complete lettering should be appended.

Authors ordinarily receive galley proofs only. Stylistic changes must be avoided in proof; excessive corrections will be charged to the authors. Authors receive gratis, postage prepaid, one hundred reprints of each article; joint authors receive fifty reprints each.

Suggestions to Authors

Footnotes should be reduced to a minimum or replaced whenever possible by a bibliography at the end of the paper; formulae in footnotes should be avoided. Authors are requested to keep in mind the typographical difficulties of complicated formulae. For example, "exp" may be used to avoid complicated exponents and a power to replace the troublesome radical sign. Symbols set one on top of another present a typographical difficulty which in most cases can easily be avoided by a stepwise arrangement.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879.

Made in United States of America